

POLYPEPTIDES THAT BIND HIV gp120 AND RELATED NUCLEIC ACIDS, ANTIBODIES, COMPOSITIONS, AND METHODS OF USE

TECHNICAL FIELD OF THE INVENTION

5       The present invention relates to polypeptides with homology to regions of domains of the human chemokine receptors CCR5, CXCR4, and STRL33, as well as domains of CD4 that bind with human immunodeficiency virus (HIV), in particular HIV-1 glycoprotein 120 (gp120) envelope protein. The present invention also relates to nucleic acids encoding such polypeptides, antibodies, compositions comprising such polypeptides, nucleic acids or antibodies, and methods of using the same.

15                      BACKGROUND OF THE INVENTION

There are seven transmembrane chemokine receptors that act as cofactors for HIV infection. The cofactors enable entry of HIV-1 into CD4<sup>+</sup> T cells and macrophages (Premack et al., *Nature Medicine* 2: 1174-78 (1996); and 20 Zhang et al., *Nature* 383: 768 (1996)).

The presence of chemokines has an inhibitory effect on HIV-1 attachment to, and infection of, susceptible cells. Additionally, some mutations in chemokine receptors have been shown to result in resistance to 25 HIV-1 infection. For example, a 32-nucleotide deletion within the CCR5 gene has been described in subjects who remained uninfected despite repeated exposures to HIV-1 (Huang et al., *Nature Medicine* 2: 1240-43 (1996)).

Evidence also exists for the physical association of 30 a ternary complex between chemokine receptors, CD4, and HIV-1 gp120 envelope glycoprotein on cell membranes

(Lapham et al., *Science* 274: 602-05 (1996)). Receptor signaling and cell activation are probably not required for the anti-HIV-1 effect of chemokines since a RANTES analog lacking the first eight amino-terminal amino acids, RANTES (9-68), lacked chemotactic and leukocyte-activating properties, but bound to multiple chemokine receptors and inhibited infection by macrophage-tropic HIV-1 (Arenzana-Seladedos et al., *Nature* 383: 400 (1996)). Cumulatively, the above described results suggest that the interaction between gp120, CD4, and at least one chemokine receptor is obligatory for HIV-1 infection. Accordingly, reagents that interfere with the binding of gp120 to chemokine receptors and to CD4 are used in the biological and medical arts. However, there presently exists a need for additional reagents that can compete with one or more proteins of the gp120-CD4-chemokine receptor complex to assist in basic biological or viral research, and to assist in medical intervention in the HIV-1 pandemic. It is an object of the present invention to provide such reagents. This and other objects and advantages, including additional inventive features, will be apparent from the description provided herein.

25

#### BRIEF SUMMARY OF THE INVENTION

The present invention provides a polypeptide that binds with HIV gp120 under physiological conditions. Multiple embodiments of the present inventive polypeptide are provided, and each embodiment possesses a degree of homology to at least one of the human CCR5, CXCR4 and

STRL33 chemokine receptors, and the human CD4 cell-surface protein.

In a first embodiment, the present invention provides a polypeptide comprising the amino acid sequence 5 YDIXYYYYXE, wherein X is any synthetic or naturally occurring amino acid residue, and the polypeptide comprises less than about 100 contiguous amino acids that are identical to, or, in the alternative, substantially identical to, the amino acid sequence of the human CCR5 chemokine receptor. A preferred polypeptide of this 10 first embodiment comprises the amino acid sequence YDIN\*YYT\*S\*E. A more preferred polypeptide of this first embodiment comprises the amino acid sequence YDINYYTSE, wherein each letter is the standard one-letter 15 abbreviation for an amino acid residue (i.e., for example, N denotes asparaginyl, T denotes threoninyl, and S denotes serinyl). The polypeptide of the first embodiment can comprise the amino acid sequence M\*D\*YQ\*V\*S\*SP\*IYDIN\*YYT\*S\*E. Preferably, the polypeptide 20 comprises the amino acid sequence MDYQVSSPIYDINYYTSE.

In a second embodiment, the present invention provides a polypeptide comprising the amino acid sequence XEXIXIYXXXNYXXX, wherein X is any synthetic or naturally occurring amino acid and wherein said polypeptide 25 comprises less than about 100 contiguous amino acid that are identical to or substantially identical to the amino acid sequence of the human CXCR4 chemokine receptor. The polypeptide can consist essentially of, or consist of, the sequence EXIXIYXXXNY. Preferably, the polypeptide 30 comprises the sequence M\*EG\*IS\*IYT\*S\*D\*NYT\*E\*E\*.

Preferably, M\*EG\*IS\*IYT\*S\*D\*NYT\*E\*E\* is  
M\*EGISIYTSDNYT\*E\*E\*.

- In a third embodiment, the present invention provides a polypeptide comprising the amino acid sequence 5 EHQAFLQFS, wherein said polypeptide comprises less than about 100 contiguous amino acids that are identical to or substantially identical to the amino acid sequence of the human STRL33 chemokine receptor. The polypeptide can consist essentially of, or consist of, the sequence 10 EHQAFLQFS.

- In a fourth embodiment, the present invention provides a polypeptide comprising at least a portion of an amino acid sequence selected from the group consisting of LPPLYSLVFIFGFVGNML, QWDFGNTMCQLLTGLYFIGFFS, 15 SQYQFWKNFQTLKIVILG, APYNIVLLLNTFQEFFFGLNNCS, and YAFVGEKFRNYLLVFFQK, wherein said polypeptide comprises less than about 100 contiguous amino acids that are identical to or substantially identical to the amino acid sequence of the human CCR5 chemokine receptor.

- 20 In a fifth embodiment, the present invention provides a polypeptide comprising at least a portion of an amino acid sequence selected from the group consisting of LLLTIPDFIFANVSEADD, VVFQFHIMVGLILPGIV, and IDSFILEIIKQGCEFEN, wherein said polypeptide comprises 25 less than about 100 contiguous amino acids that are identical to or substantially identical to the amino acid sequence of the human CXCR4 chemokine receptor.

- In a sixth embodiment, the present invention provides a polypeptide comprising at least a portion of 30 an amino acid sequence selected from the group consisting of LVISIFYHKLQSLTDVFL, PFWAYAGIHEWVFGQVMC,

EAISTVVVLATQMTLGGFL, LTMIVCYSVIIKTLHAG,  
MAVFLLTQMPFNLMKFIRSTHW, HWEYYAMTSFHYTIMVTE,  
ACLNPVLYAFVSLKFRKN and SKTFSASHNVEATSMFQL, wherein said  
polypeptide comprises less than about 100 contiguous  
5 amino acids that are identical to or substantially  
identical to the amino acid sequence of the human STRL33  
chemokine receptor.

In a seventh embodiment, the present invention  
provides a polypeptide comprising at least a portion of  
10 an amino acid sequence selected from the group consisting  
of DTYICEVED, EEVQLLVFGLTANS, THLLOGQSLTLTLES, and  
GEQVEFSFPLAFTVE, wherein said polypeptide comprises less  
than about 100 contiguous amino acids that are identical  
to or substantially identical to the amino acid sequence  
15 of the human CD4 cell-surface protein.

In the fourth to seventh embodiments, any selected  
portion of the polypeptide can comprise from 1 to about 6  
conservative amino acid substitutions. In an  
alternative, the polypeptide can be partially defined by  
20 an absence of a polypeptide sequence, outside the region  
of the portion selected from the foregoing sequences,  
that has five, or ten, contiguous amino acid residues  
that have a sequence that consists of an amino acid  
sequence that is identical to or substantially identical  
25 to the protein to which the polypeptide has homology  
(i.e., CCR5, CXCR4, STRL33, or CD4). In yet another  
alternative, the polypeptide can lack a sequence of five  
or ten contiguous amino acids which are identical to or  
substantially identical to the sequence of the protein  
30 with which the sequence has homology except that one or  
more conservatively or neutrally substituted amino acids

replace part of the sequence of the protein to which the polypeptide has homology. Additionally, any embodiment of the present inventive polypeptide can also comprise a pharmaceutically acceptable substituent.

5 Any embodiment of the present inventive polypeptide can be incorporated into a composition, which further comprises a carrier. Any suitable embodiment of the present inventive polypeptide can be encoded by a nucleic acid that can be expressed in a cell. In this regard,  
10 the present invention further provides a vector comprising such a nucleic acid. The nucleic acids and vectors also can be incorporated into a composition comprising a carrier.

15 Additionally, the present invention provides a method of making an antibody to a polypeptide of the present invention. The present invention also provides a method of prophylactically or therapeutically treating an HIV infection in a mammal.

20 Additionally, the present invention provides an anti-idiotypic antibody comprising an internal image of a portion of gp120, as well as a method of selecting such an antibody.

25 The present invention also provides a method of making an antibody to a portion of the gp120 protein that binds with a portion of CCR5, CXCR4, STRL33, or CD4, as well as the immunizing compound used to make the antibody, and the antibody itself. In another embodiment of the present invention, a method of removing HIV-1 from a bodily fluid is provided.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a listing of synthetic amino acids available (from Bachem, King of Prussia, PA) for incorporation into polypeptides of the present invention.

5

## **DETAILED DESCRIPTION OF THE INVENTION**

The present invention provides a polypeptide that binds with gp120 of HIV, in particular HIV-1, more particularly HIV-1<sub>LA1</sub>, under physiological conditions.

- 10 The polypeptide has a number of uses including, but not limited to, the use of the polypeptide to elucidate the mechanism by which HIV, such as HIV-1, attaches to and/or infects a particular cell, to induce an immune response in a mammal, in particular a human, to HIV, in particular  
15 HIV-1, and to inhibit the replication of HIV, in particular HIV-1, in an infected mammal, in particular a human.

Multiple embodiments of the present inventive polypeptide are provided. Each embodiment of the polypeptide has a degree of homology to at least one of the human CCR5, CXCR4 and STRL33 chemokine receptors, or the human CD4 cell-surface protein. In each embodiment provided herein, a letter indicates the standard amino acid designated by that letter, and a letter followed directly by an asterisk (\*) preferably represents the amino acid represented by the letter (e.g., N represents asparaginyl and T represents threoninyl), or a synthetic or naturally occurring conservative or neutral substitution therefor. Additionally, in accordance with convention, all amino acid sequences provided herein are given either from left to right, or top to bottom, such

that the first amino acid is amino-terminal and the last is carboxyl-terminal. The synthesis of polypeptides, either synthetically (i.e., chemically) or biologically, is within the skill in the art.

5 It is within the skill of the ordinary artisan to  
select synthetic and naturally occurring amino acids that  
make conservative or neutral substitutions for any  
particular naturally occurring amino acids. The skilled  
artisan desirably will consider the context in which any  
10 particular amino acid substitution is made, in addition  
to considering the hydrophobicity or polarity of the  
side-chain, the general size of the side chain, and the  
pK value of side-chains with acidic or basic character  
under physiological conditions. For example, lysine,  
15 arginine, and histidine are often suitably substituted  
for each other, and more often arginine and lysine. As  
is known in the art, this is because all three amino  
acids have basic side chains, whereas the pK value for  
the side-chains of lysine and arginine are much closer to  
20 each other (about 10 and 12) than to histidine (about 6).  
Similarly, glycine, alanine, valine, leucine, and  
isoleucine are often suitably substituted for each other,  
with the proviso that glycine is frequently not suitably  
substituted for the other members of the group. This is  
25 because each of these amino acids are relatively  
hydrophobic when incorporated into a polypeptide, but  
glycine's lack of an  $\alpha$ -carbon allows the phi and psi  
angles of rotation (around the  $\alpha$ -carbon) so much  
conformational freedom that glycinyl residues can trigger  
30 changes in conformation or secondary structure that do  
not often occur when the other amino acids are

substituted for each other. Other groups of amino acids frequently suitably substituted for each other include, but are not limited to, the group consisting of glutamic and aspartic acids; the group consisting of

- 5 phenylalanine, tyrosine and tryptophan; and the group consisting of serine, threonine and, optionally, tyrosine. Additionally, the skilled artisan can readily group synthetic amino acids with naturally occurring amino acids.

- 10 In the context of the present invention, a polypeptide is "substantially identical" to another polypeptide if it comprises at least about 80% identical amino acids. Desirably, at least about 50% of the non-identical amino acids are conservative or neutral 15 substitutions. Also, desirably, the polypeptides differ in length (i.e., due to deletion mutations) by no more than about 10%.

In a first embodiment, the present invention provides a polypeptide comprising the amino acid sequence YDIXYYXXE, wherein X is any synthetic or naturally occurring amino acid residue, and the polypeptide comprises less than about 100 contiguous amino acids, preferably less than about 50 amino acids, more preferably less than about 25 amino acids, and yet more preferably less than about 13 amino acids that are identical to, or, in the alternative, substantially identical to, the amino acid sequence of the human CCR5 chemokine receptor.

- Preferably, the polypeptide of the first embodiment  
comprises YDIXYYXXE, wherein the amino moiety of the  
amino-terminal tyrosinyl residue is not bound to another

amino acid residue via a peptidic bond, and the carboxyl moiety of the glutamyl residue is not bound to another amino acid residue via a peptidic bond. However, the polypeptide can consist essentially of YDIXYYXXE and, 5 optionally, can be modified by one or more pharmaceutically acceptable substituents, such as, for example, t-boc or a saccharide.

More particularly, the polypeptide comprises the amino acid sequence YDIN\*YYT\*S\*E. Preferably, N\* is 10 asparaginyl, T\* is threoninyl, and S\* is serinyl.

The polypeptide of the first embodiment can comprise a dodecapeptide selected from the amino acid sequence M\*D\*YQ\*V\*S\*SP\*IYDIN\*YYT\*S\*E. More preferably, the polypeptide of the first embodiment comprises the amino 15 acid sequence MDYQVSSPIYDINYYTSE.

In a second embodiment, the present invention provides a polypeptide comprising the amino acid sequence XEXIXIYXXXNYXXX, wherein X is any synthetic or naturally occurring amino acid, and the polypeptide comprises less than about 100 contiguous amino acids, preferably less 20 than about 50 amino acids, and more preferably less than about 25 amino acids, that are identical to or substantially identical to the amino acid sequence of the human CXCR4 chemokine receptor. Optionally, the polypeptide consists essentially of, or consists of, the 25 sequence EXIXIYXXXNY.

In a preferred polypeptide of this second embodiment, the polypeptide comprises the amino acid sequence M\*EG\*IS\*IYT\*S\*D\*NYT\*E\*E\*. Preferably, 30 M\*EG\*IS\*IYT\*S\*D\*NYT\*E\*E\* is M\*EGISIYTSNDNYT\*E\*E\*.

In a third embodiment, the present invention provides a polypeptide comprising the amino acid sequence EHQAFLQFS, wherein the polypeptide comprises less than about 100 contiguous amino acid residues, preferably less than about 50 contiguous amino acid residues, more preferably less than about 25 contiguous amino acid residues, that are identical to or substantially identical to the amino acid sequence of the human STRL33 chemokine receptor. The polypeptide can consist essentially of, or consist of, the sequence EHQAFLQFS.

The first three embodiments of the present invention provide, among other things, polypeptides having substantial identity or identity to the amino-terminal regions of the chemokine receptors CCR5, CXCR4, and STRL33. These first three embodiments form a first group of embodiments of the present invention. The present invention also provides, in a second group of embodiments, polypeptides having substantial identity or identity to an internal region of the human chemokine receptors CCR5, CXCR4, and STRL33, as well as to the leukocyte cell-surface protein CD4.

This second group of embodiments provides a polypeptide that binds with HIV gp120 under physiological conditions and comprises at least a portion of or all of an amino acid sequence selected from the group consisting of LPPLYSLVFIFGFVGNML, QWDFGNTMCQLLTGLYFIGFFS, SQYQFWKNFQTLKIVILG, APYNIVLLLNTFOEFFGLNNCS, and YAFVGEEKFRNYLLVFFQK, wherein the polypeptide comprises less than about 100 amino acids that are identical to or substantially identical to the amino acid sequence of the human CCR5 chemokine receptor; or selected from the group

consisting of LLLTIPDFIFANVSEADD (165-182),  
VVFQFQHIMVGLLPGIV (197-214), and IDSFILEIIKQGCEFEN  
(261-278), wherein the polypeptide comprises less than  
about 100 amino acids that are identical to or  
5 substantially identical to the amino acid sequence of the  
human CXCR4 chemokine receptor; or  
selected from the group consisting of  
LVISIFYHKLQLSLTDVFL (53-70), PFWAYAGIHEWVFGQVMC (85-102),  
EAISTVVLATQMTLGFFL (185-202), LTMIVCYSVIKTLLHAG (205-  
10 222), MAVFLLTQMPFNLMKFIRSTHW (237-258),  
HWEYYAMTSFHYTIMVTE (257-274), ACLNPVLYAFVSLKFRKN (281-  
298) and SKTFSASHNVEATSMFQL (325-342), wherein the  
polypeptide comprises less than about 100 amino acids  
that are identical to a substantially identical to the  
15 amino acid sequence of the human STRL33 chemokine  
receptor; or  
selected from the group consisting of DTYICEVED,  
EEVQLLVFGLTANS, THLLQQQLTLTLES, and GEQVEFSFPLAFTVE,  
wherein the polypeptide binds with HIV gp120 under  
20 physiological conditions and comprises less than about  
100 amino acids that are identical to or substantially  
identical to the amino acid sequence of the human CD4  
cell-surface protein. Optionally, the recited amino acid  
sequences can comprise 1 to about 6 conservative or  
25 neutral amino acid substitutions.

The polypeptides of this second group of embodiments  
preferably comprise less than about 50 amino acid  
residues, and more preferably less than about 25 amino  
acid residues, and yet more preferably no additional  
30 amino acid residues, that are identical to a protein that  
naturally has the recited amino acid sequence. The

polypeptide can be alternatively characterized by an absence of a region, outside the above-recited amino acid sequences, that has about five, or about ten, contiguous amino acid residues that have a sequence that consists of

- 5 an amino identical and conservatively substituted residues as an amino acid sequence of the protein to which the polypeptide of the compound has homology.

Any embodiment of the present inventive polypeptide can also comprise a pharmaceutically acceptable

- 10 substituent, attachment of which is within the skill in  
the art. The pharmaceutically acceptability of  
substituents are understood by those skilled in the art.  
For example, a pharmaceutically acceptable substituent  
can be a biopolymer, such as a polypeptide, an RNA, a  
15 DNA, or a polysaccharide. Suitable polypeptides comprise  
fusion proteins, an antibody or fragment thereof, a cell  
adhesion molecule or a fragment thereof, or a peptide  
hormone. Suitable polysaccharides comprise polyglucosidic  
moieties, such as starch and their derivatives, such as  
20 heparin. The pharmaceutically acceptable substituent  
also can be any suitable lipid or lipid-containing  
moiety, such as a lipid of a liposome or a vesicle, or  
even a lipophilic moiety, such as a prostaglandin, a  
steroid hormone, or a derivative thereof. Additionally,  
25 the pharmaceutically acceptable substituent can be a  
nucleotide or nucleoside, such as nicotine adenine  
dinucleotide or thymine, an amino acid residue, a  
saccharide or disaccharide, or the residue of another  
biomolecule naturally occurring in a cell, such as  
30 inositol, a vitamin, such as vitamin C, thiamine, or  
nicotinic acid. Synthetic organic moieties also can be

pharmaceutically acceptable substituents, such as t-butyl carbonyl, an acetyl moiety, quinine, polystyrene and other biologically acceptable polymers. Optionally, a pharmaceutically acceptable substituent can be selected from the group consisting of a C<sub>1</sub>-C<sub>18</sub> alkyl, a C<sub>2</sub>-C<sub>18</sub> alkenyl, a C<sub>2</sub>-C<sub>18</sub> alkynyl, a C<sub>6</sub>-C<sub>18</sub> aryl, a C<sub>7</sub>-C<sub>18</sub> alkaryl, a C<sub>7</sub>-C<sub>18</sub> aralkyl, and a C<sub>3</sub>-C<sub>18</sub> cycloalkyl, wherein any of the foregoing moieties that are cyclic comprise from 0 to 2 atoms per carbocyclic ring, which can be the same or different, and are selected from the group consisting of nitrogen, oxygen, and sulfur.

Any of the substituents from this group can be substituted by one to six substituent moieties, which can be the same or different, selected from the group consisting of an amino moiety, a carbamate moiety, a carbonate moiety, hydroxyl, a phosphamate moiety, a phosphate moiety, a phosphonate moiety, a pyrophosphate moiety, a triphosphate moiety, a sulfamate moiety, a sulfate moiety, a sulfonate moiety, a C<sub>1</sub>-C<sub>8</sub> monoalkylamine moiety, a C<sub>1</sub>-C<sub>8</sub> dialkylamine moiety, and a C<sub>1</sub>-C<sub>8</sub> trialkylamine moiety.

Any embodiment of the present inventive polypeptide can be encoded by a nucleic acid and can be expressed in a cell. The skilled artisan will recognize that the encoded polypeptide as well as any pharmaceutically acceptable substituent to be incorporated into the polypeptide, e.g., a formyl or acetyl substituent on an amino-terminal methionine or a saccharide, will preferably be produced by a cell that can express the polypeptide of the present invention. Accordingly, the

amino acids incorporated into the polypeptide encoded by the nucleic acid are preferably naturally occurring.

- A nucleic acid as described above can be cloned into any suitable vector and can be used to transduce, transform, or transfect any suitable host. The selection of vectors and methods to construct them are commonly known to persons of ordinary skill in the art and are described in general technical references (see, in general, "Recombinant DNA Part D," *Methods in Enzymology*, Vol. 153, Wu and Grossman, eds., Academic Press (1987)).
- Desirably, the vector comprises regulatory sequences, such as transcription and translation initiation and termination codons, which are specific to the type of host (e.g., bacterium, fungus, plant, or animal) into which the vector is to be inserted, as appropriate and taking into consideration whether the vector is DNA or RNA. Preferably, the vector comprises regulatory sequences that are specific to the genus of the host. Most preferably, the vector comprises regulatory sequences that are specific to the species of the host and is optimized for the expression of an above-described polypeptide.

Constructs of vectors, which are circular or linear, can be prepared to contain an entire nucleic acid sequence as described above or a portion thereof ligated to a replication system that is functional in a prokaryotic or eukaryotic host cell. Replication systems can be derived from ColE1, 2  $\mu$  plasmid,  $\lambda$ , SV40, bovine papilloma virus, and the like.

Suitable vectors include those designed for propagation and expansion, or for expression, or both. A

preferred cloning vector is selected from the group consisting of the pUC series, the pBluescript series (Stratagene, LaJolla, CA), the pET series (Novagen, Madison, WI), the pGEX series (Pharmacia Biotech, 5 Uppsala, Sweden), and the pEX series (Clonetech, Palo Alto, CA). Examples of animal expression vectors include PEUK-C1, pMAM and pMAMneo (Clonetech, Palo Alto, CA).

An expression vector can comprise a native or nonnative promoter operably linked to a nucleic acid 10 molecule encoding an above-described polypeptide. The selection of promoters, e.g., strong, weak, inducible, tissue-specific and developmental-specific, is within the skill in the art. Similarly, the combining of a nucleic acid molecule as described above with a promoter is also 15 within the skill in the art.

The skilled artisan will also recognize that the polypeptide has ability to bind the gp120 protein, which is most often found outside of cells. Accordingly, the present inventive nucleic acid advantageously can 20 comprise a nucleic acid sequence that encodes a signal sequence such that a signal sequence is translated as a fusion protein with the polypeptide of the present inventive polypeptide to form a signal sequence-polypeptide fusion. The signal sequence can cause 25 secretion of the entire polypeptide, including the signal sequence (which is a pharmaceutically acceptable substituent), or can be cleaved from the polypeptide (i.e., the polypeptide of the compound) prior to, or during, secretion so that at least the present inventive 30 polypeptide is secreted out of a cell in which the nucleic acid is expressed.

Alternatively, the nucleic acid comprises or encodes an antisense nucleic acid molecule or a ribozyme that is specific for a specified amino acid sequence of an above-described polypeptide. A nucleic acid sequence 5 introduced in antisense suppression generally is substantially identical to at least a portion of the endogenous gene or gene to be repressed, but need not be identical. Thus, the vectors can be designed such that the inhibitory effect applies to other proteins within a 10 family of genes exhibiting homology or substantial homology to the target gene. The introduced sequence also need not be full-length relative to either of the primary transcription product or the fully processed mRNA. Generally, higher homology can be used to 15 compensate for the use of a shorter sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and homology of non-coding segments will be equally effective.

Ribozymes also have been reported to have use as a 20 means to inhibit expression of endogenous genes. It is possible to design ribozymes that specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. In carrying out this 25 cleavage, the ribozyme is not itself altered and is, thus, capable of recycling and cleaving other molecules, making it a true enzyme. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of 30 the constructs. The design and use of target RNA-

specific ribozymes is described in Haseloff et al., *Nature* 334: 585-591 (1988).

Further provided by the present invention is a composition comprising an above-described polypeptide or nucleic acid and a carrier therefor. Another composition provided by the present invention is a composition comprising an antibody to an above-described polypeptide or an anti-antibody to an above-described polypeptide.

Any embodiment of the present invention including the present inventive polypeptide, nucleic acid, antibody, and anti-antibody, can be incorporated into a composition comprising a carrier. The carrier can serve any function. For example, the carrier can increase the solubility of the present inventive polypeptide, nucleic acid or antibody in aqueous solutions. Additionally, the carrier can protect the present inventive polypeptide, nucleic acid or antibody from environmental insults, such as dehydration, oxidation, and photolysis. Moreover, the carrier can serve as an adjuvant, or as a timed-release control means in a biological system.

Antibodies can be generated in accordance with methods known in the art. See, for example, Benjamin, *In Immunology: a short course*, Wiley-Liss, NY, 1996, pp. 436-437; Kuby, *In Immunology*, 3rd. ed., Freeman, NY, 1997, pp. 455-456; Greenspan et al., *FASEB J.* 7: 437-443 (1993); and Poskitt, *Vaccine* 9: 792-796 (1991). Anti-antibodies (i.e., anti-idiotypic antibodies) also can be generated in accordance with methods known in the art (see, for example, Benjamin, *In Immunology: a short course*, Wiley-Liss, NY, 1996, pp. 436-437; Kuby, *In Immunology*, 3rd. ed., Freeman, NY, 1997, pp. 455-456;

Greenspan et al., FASEB J., 7, 437-443, 1993; Poskitt, Vaccine, 9, 792-796, 1991; and Madiyalakan et al., Hybridonor 14: 199-203 (1995) ("Anti-idiotype induction therapy"). Such antibodies can be obtained and employed either in solution-phase or coupled to a desired solid-phase matrix. Having in hand such antibodies, one skilled in the art will further appreciate that such antibodies, using well-established procedures (e.g., such as described by Harlow and Lane (1988, supra), are useful in the detection, quantification, or purification of gp120 or HIV, particularly HIV-1, conjugates of each and host cells transformed to produce a gp120 receptor or a derivative thereof. Such antibodies are also useful in a method of prevention or treatment of a viral infection and in a method of inducing an immune response to HIV as provided herein.

In view of the above, an above-described polypeptide can be administered to an animal. The animal generates anti-polypeptide antibodies. Among the anti-polypeptide antibodies generated or induced in the animal are antibodies that have an internal image of gp120. In accordance with well-known methods, polyclonal or monoclonal antibodies can be obtained, isolated and selected. Selection of an anti-polypeptide antibody that has an internal image of gp120 can be based upon competition between the anti-polypeptide antibody and gp120 for binding to an above-described polypeptide, or upon the ability of the anti-polypeptide antibody to bind to a free polypeptide as opposed to a polypeptide bound to gp120. Such an anti-antibody can be administered to

an animal to prevent or treat an HIV infection in accordance with methods provided herein.

Although nonhuman anti-idiotypic antibodies, such as an anti-polypeptide antibody that has an internal image of gp120 and, therefore, is anti-idiotypic to gp120, are useful for prophylaxis in humans, their favorable properties might, in certain instances, can be further enhanced and/or their adverse properties further diminished, through "humanization" strategies, such as those recently reviewed by Vaughan, Nature Biotech., 16, 535-539, 1998.

Prior to administration to an animal, such as a mammal, in particular a human, an above-described polypeptide, nucleic acid, antibody or anti-antibody can be formulated into various compositions by combination with appropriate carriers, in particular, pharmaceutically acceptable carriers or diluents, and can be formulated to be appropriate for either human or veterinary applications.

The present invention also provides a method of making an antibody. The method comprises administering an immunogenic amount of an above-described polypeptide or nucleic acid to an animal, such as a mammal, in particular a human. Determining the quantity of a polypeptide or nucleic acid that is immunogenic will depend in part on the degree of similarity to a protein or other molecule of the inoculated animal, the route of administration of the polypeptide or nucleic acid, and the size of the polypeptide administered or encoded by the administered nucleic acid. If necessary, the polypeptide or nucleic acid can be mixed with or ligated

to a substance (or an adjuvant) that enhances its immunogenicity. Such calculations and procedures are within the skill of the ordinary artisan. Additionally, the present inventive method preferably can be used to induce an immune response against HIV, particularly HIV-1, in a mammal, particularly a human.

In view of the above, the present invention further provides a method of prophylactically or therapeutically treating an HIV infection in a mammal, particularly a human, in need thereof. The method comprises administering to the mammal an HIV replication-inhibiting effective amount of an above-described polypeptide, nucleic acid, or an anti-antibody to an above-described polypeptide or a nucleic acid encoding such a polypeptide.

The present invention also provides a method of prophylactically or therapeutically treating HIV infection in a mammal. The method comprises administering to the mammal an effective amount of an above-described polypeptide or nucleic acid. Prior to administration to an animal, such as a mammal, in particular a human, an above-described polypeptide or nucleic acid can be formulated into various compositions by combination with appropriate carriers, in particular, pharmaceutically acceptable carriers or diluents, and can be formulated to be appropriate for either human or veterinary applications.

Thus, a composition for use in the method of the present invention can comprise one or more of the polypeptides, nucleic acids, antibodies or anti-antibodies described herein, preferably in combination

with a pharmaceutically acceptable carrier.

Pharmaceutically acceptable carriers are well-known to those skilled in the art, as are suitable methods of administration. The choice of carrier will be determined, in part, by whether a polypeptide or a nucleic acid is to be administered, as well as by the particular method used to administer the composition. Optionally, the carrier can be selected to increase the solubility of the composition or mixture, e.g., a liposome or polysaccharide. One skilled in the art will also appreciate that various routes of administering a composition are available, and, although more than one route can be used for administration, a particular route can provide a more immediate and more effective reaction than another route. Accordingly, there are a wide variety of suitable formulations of compositions that can be used in the present inventive methods.

A composition in accordance with the present invention, alone or in further combination with one or more other active agents, can be made into a formulation suitable for parenteral administration, preferably intraperitoneal administration. Such a formulation can include aqueous and nonaqueous, isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and nonaqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The formulations can be presented in unit dose or multi-dose sealed containers, such as ampules and vials, and can be

stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid carrier, for example, water, for injections, immediately prior to use. Extemporaneously injectable solutions and suspensions can be prepared from sterile powders, granules, and tablets, as described herein.

A formulation suitable for oral administration can consist of liquid solutions, such as an effective amount of the compound dissolved in diluents, such as water, saline, or fruit juice; capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as solid or granules; solutions or suspensions in an aqueous liquid; and oil-in-water emulsions or water-in-oil emulsions. Tablet forms can include one or more of lactose, mannitol, corn starch, potato starch, microcrystalline cellulose, acacia, gelatin, colloidal silicon dioxide, croscarmellose sodium, talc, magnesium stearate, stearic acid, and other excipients, colorants, diluents, buffering agents, moistening agents, preservatives, flavoring agents, and pharmacologically compatible carriers.

Similarly, a formulation suitable for oral administration can include lozenge forms, which can comprise the active ingredient in a flavor, usually sucrose and acacia or tragacanth; pastilles comprising the active ingredient in an inert base, such as gelatin and glycerin, or sucrose and acacia; and mouthwashes comprising the active ingredient in a suitable liquid carrier; as well as creams, emulsions, gels, and the like containing, in addition to the active ingredient, such carriers as are known in the art.

TECHNICAL DATA SHEET

An aerosol formulation suitable for administration via inhalation also can be made. The aerosol formulation can be placed into a pressurized acceptable propellant, such as dichlorodifluoromethane, propane, nitrogen, and 5 the like.

A formulation suitable for topical application can be in the form of creams, ointments, or lotions.

A formulation for rectal administration can be presented as a suppository with a suitable base 10 comprising, for example, cocoa butter or a salicylate. A formulation suitable for vaginal administration can be presented as a pessary, tampon, cream, gel, paste, foam, or spray formula containing, in addition to the active 15 ingredient, such carriers as are known in the art to be appropriate.

Important general considerations for design of delivery systems and compositions, and for routes of administration, for polypeptide drugs also apply 20 (Eppstein, CRC Crit. Rev. Therapeutic Drug Carrier Systems 5, 99-139, 1988; Siddiqui et al., CRC Crit. Rev. Therapeutic Drug Carrier Systems 3, 195-208, 1987); Banga et al., Int. J. Pharmaceutics 48, 15-50, 1988; Sanders, Eur. J. Drug Metab. Pharmacokinetics 15, 95-102, 1990; Verhoef, Eur. J. Drug Metab. Pharmacokinetics 15, 83-93, 25 1990). The appropriate delivery system for a given polypeptide will depend upon its particular nature, the particular clinical application, and the site of drug action. As with any protein drug, oral delivery will likely present special problems, due primarily to 30 instability in the gastrointestinal tract and poor absorption and bioavailability of intact, bioactive drug

therefrom. Therefore, especially in the case of oral delivery, but also possibly in conjunction with other routes of delivery, it will be necessary to use an absorption-enhancing agent in combination with a given 5 polypeptide. A wide variety of absorption-enhancing agents have been investigated and/or applied in combination with protein drugs for oral delivery and for delivery by other routes (Verhoef, 1990, supra; van Hoogdalem, Pharmac. Ther. 44, 407-43, 1989; Davis, J. 10 Pharm. Pharmacol. 44(Suppl. 1), 186-90, 1992). Most commonly, typical enhancers fall into the general categories of (a) chelators, such as EDTA, salicylates, and N-acyl derivatives of collagen, (b) surfactants, such as lauryl sulfate and polyoxyethylene-9-lauryl ether, (c) 15 bile salts, such as glycholate and taurocholate, and derivatives, such as taurodihydrofusidate, (d) fatty acids, such as oleic acid and capric acid, and their derivatives, such as acylcarnitines, monoglycerides, and diglycerides, (e) non-surfactants, such as unsaturated 20 cyclic ureas, (f) saponins, (g) cyclodextrins, and (h) phospholipids.

Other approaches to enhancing oral delivery of protein drugs can include the aforementioned chemical modifications to enhance stability to gastrointestinal 25 enzymes and/or increased lipophilicity. Alternatively, the protein drug can be administered in combination with other drugs or substances that directly inhibit proteases and/or other potential sources of enzymatic degradation of proteins. Yet another alternative approach to prevent 30 or delay gastrointestinal absorption of protein drugs is to incorporate them into a delivery system that is

designed to protect the protein from contact with the proteolytic enzymes in the intestinal lumen and to release the intact protein only upon reaching an area favorable for its absorption. A more specific example of 5 this strategy is the use of biodegradable microcapsules or microspheres, both to protect vulnerable drugs from degradation, as well as to effect a prolonged release of active drug (Deasy, in Microencapsulation and Related Processes, Swarbrick, ed., Marcell Dekker, Inc.: New York, 1984, pp. 1-60, 88-89, 208-11). Microcapsules also 10 can provide a useful way to effect a prolonged delivery of a protein drug after injection (Maulding, J. Controlled Release 6, 167-76, 1987).

The dose administered to an animal, such as a 15 mammal, particularly a human, in the context of the present invention should be sufficient to effect a therapeutic or prophylactic response in the individual over a reasonable time frame. The dose will be determined by the particular polypeptide, nucleic acid, 20 antibody, or anti-antibody administered, the severity of any existing disease state, as well as the body weight and age of the individual. The size of the dose also will be determined by the existence of any adverse side effects that may accompany the use of the particular 25 polypeptide, nucleic acid, antibody or anti-antibody employed. It is always desirable, whenever possible, to keep adverse side effects to a minimum.

The dosage can be in unit dosage form, such as a tablet or capsule. The term "unit dosage form" as used 30 herein refers to physically discrete units suitable as unitary dosages for human and animal subjects, each unit

containing a predetermined quantity of a vector, alone or in combination with other active agents, calculated in an amount sufficient to produce the desired effect in association with a pharmaceutically acceptable diluent,

5 carrier, or vehicle. The specifications for the unit dosage forms of the present invention depend on the particular embodiment employed and the effect to be achieved, as well as the pharmacodynamics associated with each polypeptide, nucleic acid or anti-antibody in the

10 host. The dose administered should be an "HIV infection inhibiting amount" of an above-described polypeptide or nucleic acid or an "immune response-inducing effective amount" of an above-described polypeptide, an above-described nucleic acid, or an antibody as appropriate.

15 Another composition provided by the present invention is a composition comprising a solid support matrix to which is attached an above-described polypeptide, or an anti-antibody to an above-described polypeptide. The solid matrix can comprise other

20 functional reagents including, for example, polyethylene glycol, dextran, albumin and the like, whose intended effector functions may include one or more of the following: to improve stability of the conjugate; to increase the half-life of the conjugate; to increase resistance of the conjugate to proteolysis; to decrease the immunogenicity of the conjugate; to provide a means to attach or immobilize a functional polypeptide or anti-antibody onto a solid support matrix (e.g., see, for example, Harris, in Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications, Harris, ed.,

25 Plenum Press: New York (1992), pp. 1-14). Conjugates

SEARCHED  
INDEXED  
SERIALIZED  
FILED

furthermore may comprise a polypeptide or anti-antibody coupled to an effector molecule, each of which, optionally, may have different functions (e.g., such as a toxin molecule (or an immunological reagent) and a 5 polyethylene glycol (or dextran or albumin) molecule). Diverse applications and uses of functional proteins and polypeptides, attached to or immobilized on a solid support matrix, are exemplified more specifically for poly(ethylene glycol) conjugated proteins or peptides in 10 a review by Holmberg et al. (In Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications, Harris, ed., Plenum Press: New York, 1992, pp. 303-324).

In addition, the present invention provides a method of removing HIV from a bodily fluid of an animal. The 15 method comprises extracorporeally contacting the bodily fluid of the animal with a solid-support matrix to which is attached an above-described polypeptide or an anti-antibody to an above-described polypeptide. Alternatively, the bodily fluid can be contacted with the 20 polypeptide or anti-antibody in solution and then the solution can be contacted with a solid support matrix to which is attached a means to remove the polypeptide or anti-antibody to which is bound HIV gp120 from the bodily fluid.

25 Methods of attaching an herein-described polypeptide, or an anti-antibody to a solid support matrix are known in the art. "Attached" is used herein to refer to attachment to (or coupling to) and immobilization in or on a solid support matrix. See, for example, Harris, in Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications, Harris, ed., 30

RECEIVED  
MAY 22 1996  
U.S. PATENT AND TRADEMARK OFFICE

- Plenum Press: New York (1992), pp. 1-14) and international patent application WO 91/02714 (Saxinger). Diverse applications and uses of functional polypeptides attached to or immobilized on a solid support matrix are exemplified more specifically for poly(ethylene glycol) conjugated proteins or peptides in a review by Holmberg et al. (In Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications, Harris, ed., Plenum Press: New York, 1992, pp. 303-324).
- The present invention also provides a method of making an antibody that binds to gp120 of HIV under physiological conditions. The method comprises labeling an embodiment of the present inventive compound to obtain a labeled compound. Labeling compounds are within the skill of the ordinary artisan. For example, the present inventive compound can be labeled with radioactive atom, such as <sup>125</sup>I in the same or a similar manner as was performed in the examples provided below. Alternatively, an enzyme, such as horseradish peroxidase, can be attached to or incorporated into the present inventive compound. Then by exposing a chromogenic or photogenic compound to the compound, a signal indicative of the presence and quantity of the compound present can be generated. In another alternative, a polyhistidinyl moiety can be attached to, or incorporated into, the present inventive moiety so that the present inventive compound will react with high affinity to transition metal ions such as nickel, copper, or zinc ions; this reaction can be used as the basis to quantify the amount of the present inventive compound present at a particular location. In yet another alternative, the present

2009-04-22 14:52:25

inventive compound can be used as antigen to a standard antibody that specifically recognizes an antigenic epitope of the present inventive compound. As is well-known, the standard antibody can itself be labeled or 5 used in conjunction with an additional antibody that is labeled with an enzyme, radioisotope, or other suitable means. The skilled artisan will recognize that there is a plethora of other suitable means and methods to label the present inventive compound.

10 This present inventive method of making an antibody that binds to a gp120 envelope protein of HIV further comprises providing a library of synthetic peptides. The library consists of a multiplicity of synthetically-produced polypeptides that are homologous, and preferably 15 essentially identical (i.e., having the same primary amino acid residue sequence, ignoring blocking groups, phosphorylation of serinyl, threoninyl, and tyrosinyl residues, hydroxylation of prolinyl residues, and the like) or identical, to a continuous region of an HIV 20 gp120 envelope protein. The polypeptides of the library can be any suitable length. While larger regions allow faster scanning and tend to preserve non-linear epitopes, shorter length polypeptides allow more sensitive screening of the primary sequence of the gp120 protein. 25 However, polypeptides that are too short can lose essential secondary structure or cleave reactive sites into one or more pieces. Preferably, a mixture of short and long polypeptides are incorporated into the library, however, the library can consist of polypeptides of a 30 single length (measured in amino acid residues). For the sake of convenience the library can be split into

multiple parts, and screened by parts. Typically, the polypeptides of the library will be between about 6 and about 45 amino acid residues in length.

Typically, the library will comprise a series of 5 polypeptides each having an identical sequence to that of gp120 but having an amino-terminus a particular number of amino acids downstream of the amino-terminus of the prior polypeptide (see, examples section below). The distance, measured in amino acid residues, is referred to as the 10 offset. Preferably, libraries that are characterized by the existence of an offset, the offset is not greater than the product of length of the longest polypeptide measured in amino acid residues and 1.5, preferably 1.0, and more preferably 0.5. The library can be 15 alternatively characterized by the existence of an offset not greater than 30, preferably 15, and more preferably 4.

Each polypeptide of the library is substantially isolated from every other polypeptide of said library and 20 is located in a known position. For example, each polypeptide can be bound to a solid support and that is in a vessel or that can be placed in a vessel. The vessel preferably enables each polypeptide to be covered in a liquid that does not contact any other 25 oligonucleotide of the library. By way of example, each polypeptide can be bound to a bead that is placed in a vessel (or tube) or can be bound to the well of a multi-well assay plate. Alternatively, an array of polypeptides can be fashioned, for example on a microchip 30 device (as is presently used in some DNA sequencing

devices and methods), and the entire array can be bathed in a single solution.

Each polypeptide is then individually contacted with the labeled compound such that a portion of the labeled compound can bind with the polypeptide of the library.

In this way, a bound population of each labeled compound of the present invention and an unbound population of the labeled compound is generated. The phrase individually contacted means that each polypeptide has the opportunity to bind with the labeled compound and the quantity of labeled compound bound by each can be determined.

The method then comprises removing substantially all of the unbound labeled compound from the position occupied by each polypeptide. That is, the solution comprising the labeled compound is separated from the polypeptides of the library and the bound population of the labeled compound. This can be done by any suitable method, e.g., by aspiration and one or more washing steps comprising adding a quantity of liquid sufficient to cover all the surfaces that were contacted by the labeled compound and aspirating away substantially all of the wash liquid.

The amount of labeled compound that remains co-localized with each polypeptide of the library is then measured to determine the quantity of labeled compound bound by each polypeptide. The amount of the present inventive compound bound by each polypeptide can be directly evaluated to identify a portion of the HIV gp120 envelope protein that binds to an (HIV)-receptor selected from the group consisting of CCR5, CXCR4, STRL33, and CD4. This information is then used to identify and

provide an immunizing compound. The immunizing compound comprises a polypeptide comprising an amino acid sequence that is homologous to, or preferably is essentially identical to, or identical to, the portion of the HIV-1 gp120 envelope protein that binds with CD4, CCR5, CXCR4, and/or STRL33. The immunizing protein can be provided by processing gp120, e.g., proteolytically digesting gp120 that has been isolated from a preparation of HIV-1. Preferably, however, the immunizing compound is prepared synthetically, or by genetic engineering, or by a combination of genetic engineering and synthetic methods. The immunizing compound can comprise a pharmaceutically acceptable substituent, can be encoded by a nucleic acid that can be expressed in a cell, can be mixed with a carrier, and is an inventive aspect of the present invention.

An immunogenic quantity of the immunizing compound is then inserted into an animal (e.g., a human, or a rodent, a canine, a feline, or a ruminant) in a manner consistent with the discussion of a method of raising an antibody to the present inventive compounds that are homologous to portions of CCR5, CXCR4, STRL33, and CD4, above. The insertion of the immunizing compound causes the inoculated animal to produce an antibody that binds with said portion of the HIV gp120 envelope protein. Thus the present invention also provides an antibody that binds to an HIV gp120 envelope protein, as well as an antigen binding protein comprising one or more complementarity determining regions of the antibody (e.g., a Fab, a Fab<sub>2</sub>, an Fv, a single-chain antibody, a

202520-57678057

diabody, and humanized variants of all of the above, all of which are within the skill in the art).

The antibody or variant thereof is preferably useful in detecting or diagnosing the presence of HIV gp120 envelope protein, and thus HIV, in an animal. The antibody is also preferably prevents or attenuates infection of an animal exposed to HIV, to whom an effective quantity of the antibody or a variant thereof, has been administered or produced in response to inoculation with the immunizing compound. The antibody preferably also is useful in treating or preventing (i.e., inhibiting) HIV infection in an animal to whom a suitable dose has been administered or in which a suitable quantity of antibody has been produced. The antibody is also useful in the study of HIV infection of mammalian cells, the host range specificities of HIV infection, and preferably, the mechanism by which antibodies neutralize infectious viruses.

#### EXAMPLES

The following examples further illustrate the present invention but, of course, should not be construed as limiting the scope of the claimed invention in any way.

Synthetic peptide arrays were constructed in 96-well microtiter plates in accordance with the method set forth in WO 91/02714 (Saxinger), and used to test the binding of HIV-1<sub>LAI</sub> envelope gp120 that had been labeled with radioactive iodine (radiolabeling by standard methods). After incubating the radiolabeled gp120 in a well with each synthetic peptide, a washing step was performed to

202520-47348007

remove unbound label, and the relative level of radioactivity remaining in each well of the plate was evaluated to determine the relative affinity of each peptide for the gp120. The synthesis of the peptides and the quantity of binding between the synthetic peptides and the gp120 were found to be suitably reproducible, precise, and sensitive. Initial screening of the entire primary sequence of the chemokine and CD4 receptor molecules was taken 18 amino acid residues at a time.

10       The authenticity of the binding signals generated by this technique has been repeatedly demonstrated by showing that antibodies to CCR5 and CXCR4 are able to inhibit the binding of radiolabeled gp120 to the polypeptides derived from CCR5 and CXCR4 that show a high affinity for binding with gp120. Additionally, the accuracy of the binding assay used hereinbelow is demonstrated by Example 7.

Example 1

20       This example identifies segments of the CCR5 co-receptor that bind with gp120.

The first column in the table below indicates the number of the amino acid in the wild-type CCR5 receptor. The second column explicitly identifies the peptide sequence. The third column indicates the radioactive counts recorded in twenty minutes (i.e., the cpm x 20) after the background or non-specific counts had been subtracted. The fourth column contains an X in each row for which the listed polypeptide bound with high affinity to gp120. The fifth and final column contains an X in each row wherein the listed sequence binds with

substantial affinity but is weak in comparison to other samples, particularly adjacent samples.

SEQ SEG PEPTIDE	Counts per 20'	Peak Activity	non-Peak activity
	Average- background		
empty (control)	7		
1--18 MDYQVSSPIYDINYYTSE	735	X	
5--22 VSSPIYDINYYTSEPCQK	383	X	
9--26 IYDINYYTSEPCQKINVK	228	X	
13-30 NYYTSEPCQKINVKQIAA	6		
17-34 SEPCQKINVKQIAARLLP	44		
21-38 QKINVQIAARLLPPLYS	20		
25-42 VKQIAARLLPPLYSLVFI	18		
29-46 AARLLPPLYSLVFIFGFV	33		
33-50 LPPLYSLVFIFGFVGNML	705	X	
37-54 YSLVFIFGFVGNMLVILI	347	X	
41-58 FIFGFVGNMLVILILINC	343	X	
45-62 FVGNMLVILILINCKRLK	62		
49-66 MLVILILINCKRLKSMTD	84		
53-70 LILINCKRLKSMTDIYLL	2		
57-74 NCKRLKSMTDIYLLNLAI	25		
61-78 LKSMTDIYLLNLAIISDLF	210		
65-82 TDIYLLNLAIISDLFFLLT	38		
69-86 LLNLAIISDLFFLLTVFW	144		
73-90 AISDLFFLLTVFWAHYA	41		
77-94 LFFLLTVFWAHYAAAQW	173		
81-98 LTVPFWAHYAAAQWDFGN	306		
85- FWAHYAAAQWDFGNTMCQ	212		
89- YAAAQWDFGNTMCQLLTG	494	X	
93- QWDFGNTMCQLLTGLYFI	1019	X	
97- GNTMCQLLTGLYFIGFFS	941	X	
101- CQLLTGLYFIGFFSGIFF	489	X	
105- TGLYFIGFFSGIFFIILL	80		
109- FIGFFSGIFFIILLTIDR	76		
113- FSGIFFIILLTIDRYLAV	83		
117- FFIILLTIDRYLAVVHAV	77		
121- LLTIDRYLAVVHAVFALK	31		
125- DRYLAVVHAVFALKARTV	62		
129- AVVHAVFALKARTVTFGV	34		
133- AVFALKARTVTFGVVTSV	63		

137-	LKARTVTFGVVTSVITWV	74
141-	TVTGFVVTSVITWVVAVF	-25
145-	GVVTSVITWVAVFASLP	69
149-	SVITWVAVFASLPGIIF	46
153-	WVAVFASLPGIIFTRSQ	87
157-	VFASLPGIIFTRSQKEGL	54
161-	LPGIIFTRSQKEGLHYTC	118
165-	IFTRSQKEGLHYTCSSH	98
169-	SQKEGLHYTCSSHFPYSQ	304
173-	GLHYTCSSHFPYSQYQFW	301
177-	TCSSHFPYSQYQFWKNFQ	367
181-	HFPYSQYQFWKNFQTLKI	1008
185-	SQYQFWKNFQTLKIVILG	1572
189-	FWKNFQTLKIVILGLVLP	40
193-	FQTLKIVILGLVLPPLLVM	45
197-	KIVILGLVLPPLLVMVICY	65
201-	LGLVLPPLLVMVICYSIL	180
205-	LPPLLVMVICYSILKTLL	68
209-	VMVICYSILKTLLRCRN	-8
213-	CYSILKTLLRCRNEKKR	70
217-	ILKTLLRCRNEKKRHRAV	19
221-	LLRCRNEKKRHRAVRLIF	102
225-	RNEKKRHRAVRLIFTIMI	23
229-	KRHRAVRLIFTIMIVYFL	36
233-	AVRLIFTIMIVYFLFWAP	62
237-	IFTIMIVYFLFWAPYNIV	121
241-	MIVYFLFWAPYNIVLLL	214
245-	FLFWAPYNIVLLLNTFQE	616
249-	APYNIVLLLNTFQEFFFGL	1962
253-	IVLLLNTFQEFFFGLNNCS	X 2134
257-	LNTFQEFFFGLNNCSSSNR	X 293
261-	QEFFFGLNNCSSSNRLDQA	X 63
265-	GLNNCSSSNRLDQAMQVT	-31
269-	CSSNRLLDQAMQVTETLG	90
273-	NRLDQAMQVTETLGMTHC	10
277-	QAMQVTETLGMTHCCINP	81
281-	VTETLGMTHCCINPIYA	15
285-	LGMTHCCINPIIYAFVG	282
289-	HCCINPIIYAFVGKFRN	200
293-	NPIIYAFVGKFRNYLLV	162
297-	YAFVGKFRNYLLVFFQK	596
301-	GEKFRNYLLVFFQKHIAK	69

305-	RNYLLVFFQKHIAKRFCK
309-	LVFFQKHIAKRFCKCCSI
313-	QKHIAKRFCKCCSIFQQE
317-	AKRFCKCCSIFQQEAPER
321-	CKCCSIFQQEAPERASSV
325-	SIFQQEAPERASSVYTRS
329-	QEAPERASSVYTRSTGEQ
333-	ERASSVYTRSTGEQEISV
337-	SVYTRSTGEQEISVGL

65
76
23
64
53
100
84
84
47

These data indicate that, in addition to polypeptide sequences derived from positions 1-18 of the CCR5 receptor, the polypeptide sequences LPPLYSLVFIFGFVGNML, 5 QWDFGNTMCQLLTGLYFIGFFS, SQYQFWKNFQTLKIVILG, APYNIVLLLNTFQEFFGLNNCS, and YAFVGEKFRNYLLVFFQK comprise multiple subsequences, each which is capable of binding to HIV-1 envelope gp120.

10 Example 2

This example identifies segments of the CXCR4 co-receptor that bind with gp120.

The first column in the table below indicates the number of the amino acid in the wild-type CXCR4 receptor. 15 The second column explicitly identifies the peptide sequence. The third and fourth columns indicate the radioactive counts recorded in twenty minutes (i.e., the cpm x 20) after the background or non-specific counts had been subtracted. The fifth column contains an X in each row for which the listed polypeptide bound with high affinity to gp120. The sixth and final column contains an X in each row wherein the listed sequence binds with substantial affinity but is weak in comparison to other samples, particularly adjacent samples.

SEQ SEG	PEPTIDE	Major Activity Peak	Minor Activity Peak
	empty (control)		
1- 18	MEGISIYTSDNYTEEMGS	412	0
5--22	SIYTSNDNYTEEMGSGDYD	3003	2591
9--26	SDNYTEEMGSGDYDSMKE	483	71
13-30	TEEMGSGDYDSMKEPCFR	455	43
17-34	GSGDYDSMKEPCFREENA	453	41
21-38	YDSMKEPCFREENANFNK	384	-28
25-42	KEPCFREENANFNKIFLP	465	53
29-46	FREENANFNKIFLPTIYS	664	252
33-50	NANFNKIFLPTIYSIIFL	463	51
37-54	NKIFLPTIYSIIFLTGIV	585	173
41-58	LPTIYSIIFLTGIVGNGL	550	138
45-62	YSIIFLTGIVGNGLVILV	530	118
49-66	FLTGIVGNGLVILVMGYQ	535	123
53-70	IVGNGLVLVLMGYQKKLR	658	246
57-74	GLVILVMGYQKKLRSMTD	650	238
61-78	LVMGYQKKLRSMTDKYRL	569	157
65-82	YQKKLRSMTDKYRLHLSV	517	105
69-86	LRSMTDKYRLHLSVADLL	511	99
73-90	TDKYRLHLSVADLLFVIT	572	160
77-94	RLHLSVADLLFVITLPFW	504	92
81-98	SVADLLFVITLPFWAVDA	548	136
85-102	LLFVITLPFWAVDAVANW	665	253
89-106	ITLPFWAVDAVANWYFGN	475	63
93-110	FWAVDAVANWYFGNFLCK	542	130
97-114	DAVANWYFGNFLCKAVHV	478	66
101-118	NWYFGNFLCKAVHVIYTV	524	112
105-122	GNFLCKAVHVIYTVNLYS	508	96
109-126	CAVHVIYTVNLSSVLI	643	231
113-130	HVIYTVNLSSVLLAFI	655	243
117-134	TVNLSSVLLAFISLDR	530	118
121-138	YSSVLILAFISLDRLAI	654	242
125-142	LILAFISLDRLAIHVAT	569	157
129-146	FISLDRLAIHVATNSQR	519	107
133-150	DRYLAIHVATNSQRPRKL	503	91
137-154	AIVHATNSQRPRKLLAEK	580	168
141-158	ATNSQRPRKLLAEKVVYV	485	73
145-162	QRPRKLLAEKVVYVGWV	490	78
		539	127

149-166	KLLAEKVYVGWIPALL
153-170	EKVVYVGWIPELLTIP
157-174	YVGWIPALLTIPDFIF
161-178	WIPALLTIPDFIFANVS
165-182	LLLTI PDFIFANVSEADD
169-186	I PDFIFANVSEADDRYIC
173-190	I FANVSEADDRYICDRFY
177-194	VSEADDRYICDRFYPNDL
181-198	DDRYICDRFYPNDLWVVV
185-202	ICDRFYPNDLWVVVFQFQ
189-206	FYPNDLWVVVFQFHIMV
193-210	DLWVVVFQFHIMVGLIL
197-214	VVFQFHIMVGLILPGIV
201-218	FQHIMVGLILPGIVILSC
205-222	MVGLILPGIVILSCYCII
209-226	ILPGIVILSCYCIIISKL
213-230	IIVLSCYCIIISKLHSK
217-234	SCYCIIISKLHSKGHQK
221-238	IIISKLHSKGHQKRKAL
225-242	KLSHSKGHQKRKALKTTV
229-246	SKGHQKRKALKTTVIL
233-250	QKRKALKTTVILAFFA
237-254	ALKTTVILAFFACWL
241-258	TVILAFFACWLPPYYIG
245-262	ILAFFACWLPPYYIGISID
249-266	FACWLPPYYIGISIDSFIL
253-270	LPYYIGISIDSFILLEII
257-274	IGISIDSFILLEIIKGQGC
261-278	IDSFILLEIIKGCEFEN
265-282	ILLEIIKGCEFENTVHK
269-286	IIKGCEFENTVHKWISI
273-290	GCEFENTVHKWISITEAL
277-294	ENTVHKWISITEALAFFH
281-298	HKWISITEALAFFHCCLN
285-302	SITEALAFFHCCLNPILY
289-306	ALAFFFCLNPILYAFGLG
293-310	FHCCLNPILYAFLGAKFK
297-314	LNPILYAFLGAKFKTSQAQ
301-318	LYAFLGAKFKTSQAQHALT
305-322	LGAKFKTSQAQHALTSVSR
309-326	FKTSQAQHALTSVSRGSSL
313-330	AQHALTSVSRGSSLKILS

501	89
559	147
536	124
594	182
1418	1006
850	438
679	267
569	157
537	125
718	306
828	416
834	422
1001	589
582	170
579	167
604	192
689	277
671	259
569	157
542	130
552	140
695	283
673	261
735	323
596	184
614	202
851	439
1146	734
3884	3472
529	117
518	106
676	264
727	315
575	163
600	188
593	181
535	123
686	274
568	156
612	200
585	173
559	147

317-334 LTSVRGSSLKILSKGKR  
 321-338 SRGSSLKILSKGKGRRGHGS  
 325-342 SLKILSKGKGRRGHSSVST  
 329-346 LSKGKGRRGHSSVSTESES  
 333-350 KRGGHSSVSTESESSSFH  
 337-352 HSSVSTESESSSFHSS

595	183
581	169
697	285
597	185
579	167
515	103

These data indicate that, in addition to polypeptide sequences derived from positions 1-18 of the CXCR4 receptor, the polypeptide sequences LLLTIPDFIFANVSEADD (165-182), VVFQFQHIMVGLLPGIV (197-214), and IDSFILEIIKQGCEFEN (261-278) comprise multiple subsequences, which is capable of binding to HIV-1 envelope gp120.

5

10 Example 3

This example identifies segments of the STRL33 co-receptor that bind with gp120.

The first column in the table below indicates the number of the amino acid in the wild-type STRL33 receptor. The second column explicitly identifies the peptide sequence. The third and fourth columns indicate the radioactive counts recorded in twenty minutes (i.e., the cpm x 20) after the background or non-specific counts had been subtracted. The fifth column contains an X in each row for which the listed polypeptide bound with high affinity to gp120. The sixth and final column contains an X in each row wherein the listed sequence binds with substantial affinity but is weak in comparison to other samples, particularly adjacent samples.

15

20

<u>SEQ SEG</u>	<u>PEPTIDE</u>	Major Activity	Minor Activity
		<u>Peak</u>	<u>Peak</u>
	empty (control)	-34.5	34.5
1--18	MAEHDYHEDYGFSSFNDS	1178.5	1320.5
5--22	DYHEDYGFSSFNDSSQSEE	3357.5	3689.5
9--26	DYGFSSFNDSSQEEHQAF	8579.5	8909.5
13-30	SSFNDSQQEEHQAFLQFS	2689.5	2757.5
17-34	DSSQEEHQAFLQFSKVFL	869.5	2152.5
21-38	EEHQAFLQFSKVFLPCMY	2316.5	1819.5
25-42	AFLQFSKVFLPCMYLVVF	1421.5	1359.5
29-46	FSKVFLPCMYLVVFVCGL	534.5	633.5
33-50	FLPCMYLVVFVCGLVGNS	605.5	372.5
37-54	MYLVVVFVCGLVGNSLVLV	168.5	235.5
41-58	VFVCGLVGNSLVLVISIF	570.5	284.5
45-62	GLVGNSLVLVISIFYHKL	164.5	95.5
49-66	NSLVVISIFYHKLQLSLLT	1255.5	1378.5
53-70	LVISIFYHKLQLSLLTDVFL	1620.5	1780.5
57-74	IFYHKLQLSLLTDVFLVNLP	1275.5	1256.5
61-78	KLQLSLLTDVFLVNLPPLADL	412.5	348.5
65-82	LTDVFLVNLPPLADLVFC	233.5	336.5
69-86	FLVNLPPLADLVFCVTLPF	70.5	51.5
73-90	LPLADLVFCVTLPFWAYA	557.5	960.5
77-94	DLVFVCTLPFWAYAGIHE	1116.5	1063.5
81-98	VCTLPFWAYAGIHEWVFG	1819.5	1754.5
85-102	PFWAYAGIHEWVFGQVMC	7262.5	7537.5
89-106	YAGIHEWVFGQVMCKSLL	5911.5	6245.5
93-110	HEWVFGQVMCKSLLGIYT	3391.5	3466.5
97-114	FGQVMCKSLLGIYTINFY	1257.5	1354.5
101-118	MCKSLLGIYTINFYTSML	1505.5	1283.5
105-122	LLGIYTINFYTSMLILTC	499.5	408.5
109-126	YTINFYTSMLILTCITVD	351.5	510.5
113-130	FYTTSMLILTCITVDRFIV	744.5	907.5
117-134	MLILTCITVDRFIVVVKA	298.5	228.5
121-138	TCITVDRFIVVVKATKAY	89.5	346.5
125-142	VDRFIVVVKATKAYNQQA	103.5	53.5
129-146	IVVVKATKAYNQQAKRMT	166.5	43.5
133-150	KATKAYNQQAKRMTWGKV	701.5	568.5
137-154	AYNQQAKRMTWGKVTSLL	55.5	4.5
141-158	QAKRMTWGKVTSLLIWVI	-71.5	-31.5
145-162	MTWGKVTSLLIWVISLLV	-0.5	-26.5

149-166	KVTSLLIWVISLLVSLPQ	-39.5	-118.5	
153-170	LLIWVISLLVSLPQIIYG	42.5	75.5	
157-174	VISLLVSLPQIIYGNVFN	-60.5	-127.5	
161-178	LVS LPQIIYGNVFNLDKL	91.5	-15.5	
165-182	PQIIYGNVFNLDKLICGY	-18.5	-37.5	
169-186	YGNVFNLDKLICGYHDEA	-41.5	-20.5	
173-190	FNL DKLICGYHDEAISTV	1072.5	1078.5	X
177-194	KLICGYHDEAISTVVLAT	1363.5	1604.5	X
181-198	GYHDEAISTVVLATQM TL	754.5	1181.5	X
185-202	EAI STVVLATQM TLGFFL	3973.5	3745.5	X
189-206	TVVLATQM TLGFFLPLLT	2327.5	2389.5	X
193-210	ATQM TLGFFLPLLT MIVC	2365.5	2444.5	X
197-214	TLGFFLPLLT MIVCYSVI	2387.5	479.5	
201-218	FLPLLT MIVCYSVI KTL	1270.5	1195.5	X
205-222	LTMIVCYSVI KTL LHAG	2787.5	2654.5	
209-226	VCYSVI KTL LHAGGFQK	1334.5	1143.5	
213-230	VII KTL LHAGGFQK HRS L	961.5	682.5	
217-234	TLLHAGGFQK HRS LK II F	1041.5	999.5	
221-238	AGGFQK HRS LK II F LVMA	340.5	260.5	
225-242	QK HRS LK II F LVMA FLL	810.5	814.5	
229-246	SLK II F LVMA FLL T QMP	612.5	853.5	
233-250	I F LVMA F LL T QMP FN LM	386.5	772.5	
237-254	MAF VLL T QMP FN LM K FIR	2263.5	2842.5	X
241-258	LLT QMP FN LM K FIR STH W	2513.5	3154.5	X
245-262	M P FN LM K FIR STH WE YYA	2171.5	2182.5	X
249-266	LM K FIR STH WE YYA M TS F	934.5	949.5	
253-270	I R STH WE YYA M TS F HY TI	1571.5	1807.5	X
257-274	HWE YYA M TS F HY TI MVTE	2040.5	3065.5	X
261-278	YAM TS F HY TI MVTE A I A Y	2688.5	2359.5	X
265-282	S F HY TI MVTE A I A Y L RAC	761.5	1033.5	
269-286	T IM VTE A I A Y L RAC L N P V	140.5	272.5	
273-290	T E A I A Y L RAC L N P V L Y A F	604.5	480.5	
277-294	A Y L RAC L N P V L Y A F V S L K	1802.5	1849.5	
281-298	A C L N P V L Y A F V S L K F R K N	4173.5	4515.5	X
285-302	P V L Y A F V S L K F R K N F W K L	1859.5	2147.5	X
289-306	A F V S L K F R K N F W K L V K D I G C L P	808.5	1040.5	
293-310	L K F R K N F W K L V K D I G C L P	920.5	957.5	
297-314	K N F W K L V K D I G C L P Y L G V	143.5	82.5	
301-318	K L V K D I G C L P Y L G V S H Q W	-2.5	27.5	
305-322	D I G C L P Y L G V S H Q W K S S E	17.5	78.5	
309-326	L P Y L G V S H Q W K S S E D N S K	111.5	122.5	
313-330	G V S H Q W K S S E D N S K T F S A	208.5	306.5	

2006-07-26 09:22:02

317-334	QWKSSEDNSKTF SASHNV	464.5	533.5
321-338	SEDNSKTF SASHNVEATS	524.5	434.5
325-342	SKTF SASHNVEATSMFQL	1524.5	1239.5

X

These data indicate that, in addition to polypeptide sequences derived from positions 9-26 of the STRL33 receptor, the polypeptide sequences LVISIFYHKLQSLTDVFL (53-70), PFWAYAGIHEWVFGQVMC (85-102), EAISTVVLATQM TLGFFL (185-202), LTMIVCYSVIIKTL LHAG (205-222), MAVFLLTQMPFNL MKFIRSTHW (237-258), HWEYYAMTSFHYTIMVTE (257-274), ACLNPVLYAFVSLKFRKN (281-298) and SKTF SASHNVEATSMFQL (325-342) comprise multiple subsequences, which is capable of binding to HIV-1 envelope gp120.

#### Example 4

This example identifies segments of the human CD4 protein that bind with gp120.

The second column in the in the table below identifies the amino acid residue sequence of the polypeptide employed in the assay. The first column identifies the sequence coordinates of human CD4 that have an identical amino acid sequence. The third column indicates the number of radioactive decays (i.e., counts) that were counted, which is indicative of the affinity of the synthetic polypeptide for the gp120 protein. In the table below, polypeptides retaining more than 4,000 counts identify fragments that have a substantial capability to bind with gp120. Polypeptides retaining more than 6,000 counts have more substantial binding affinity. Polypeptides retaining at least about 10,000 counts have a substantial and strong capacity to bind to

gp120. Of course, fragments corresponding to amino acid coordinates 101-121 and 106-126 have a substantial, strong, and dominant capacity to bind to gp120.

B1 ( 1)	1-21	MNRGVPFRHLLLVLQLALLPA	3587
C1 ( 2)	6-26	PFRHLLLVLQLALLPAATQGK	4356
D1 ( 3)	11-31	LLVLQLALLPAATQGKKVVLG	1785
E1 ( 4)	16-36	LALLPAATQGKKVVLGKKGDT	1759
F1 ( 5)	21-41	AATQGKKVVLGKKGDTVELTC	1562
G1 ( 6)	26-46	KKVVLGKKGDTVELTCTASQK	1910
H1 ( 7)	31-51	GKKGDTVELTCTASQKKSIQF	1831
A2 ( 8)	36-56	TVELTCTASQKKSIQFHWKNS	1732
B2 ( 9)	41-61	CTASQKKSIQFHWKNSNQIKI	1717
C2 (10)	46-66	KKSIQFHWKNSNQIKILGNQNG	2182
D2 (11)	51-71	FHWKNSNQIKILGNQGSFLTK	1835
E2 (12)	56-76	SNQIKILGNQGSFLTKGPSKL	1487
F2 (13)	61-81	ILGNQGSFLTKGPSKLNDRAD	1467
G2 (14)	66-86	GSFLTKGPSKLNDRADSRSSL	1844
H2 (15)	71-91	KGPSKLNDRADSRSLWDQGN	1912
A3 (16)	76-96	LNDRADSRSLWDQGNFPLII	1753
B3 (17)	81-101	DSRRSLWDQGNFPLIIKNLKI	2224
C3 (18)	86-106	LWDQGNFPLIIKNLKIEDSDT	3264
D3 (19)	91-111	NFPLIIKNLKIEDSDTYICEV	11646
E3 (20)	96-116	IKNLKIEDSDTYICEVEDQKE	8439
F3 (21)	101-121	IEDSDTYICEVEDQKEEVQLL	6803
G3 (22)	106-126	TYICEVEDQKEEVQLLVFGLT	44965
H3 (23)	111-131	VEDQKEEVQLLVFGLTANSDT	36249
A4 (24)	116-136	EEVQLLVFGLTANSDTLLLQG	14171
B4 (25)	121-141	LVFGLTANSDTLLLQGQSLTL	3683
C4 (26)	126-146	TANSDTLLLQGQSLTLESP	6114
D4 (27)	131-151	THLLQGQSLTLESPPPGSSP	2552
E4 (28)	136-156	GQSLTLESPPPGSSPSVQCR	1538
F4 (29)	141-161	LTLESPPGSSPSVQCRSPRGK	1476
G4 (30)	146-166	PPGSSPSVQCRSPRGKNIQGG	1496
H4 (31)	151-171	PSVQCRSPRGKNIQGGKTL	1400
A5 (32)	156-176	RSPRGKNIQGGKTLSQLLEL	2066
B5 (33)	161-181	KNIQGGKTLSQLLELQDSGT	3078
C5 (34)	166-186	GKTLSQLLELQDSGTWTCTV	2618
D5 (35)	171-191	VSQLLELQDSGTWTCTVLQNQ	3879
E5 (36)	176-196	LQDSGTWTCTVLQNQKKVEFK	2456
F5 (37)	181-201	TWTCTVLQNQKKVEFKIDIV	4030
G5 (38)	186-206	VLQNQKKVEFKIDIVVLA	9737
H5 (39)	191-211	FQKASSIV	6313
A6 (40)	196-216	KIDIVVLA	3681

B6 (41)	201-221	VLAFKASSIVYKKEGEQVEF	3566
C6 (42)	206-226	KASSIVYKKEGEQVEFSFPLA	14347
D6 (43)	211-231	VYKKEGEQVEFSFPLAFTVEK	14740
E6 (44)	216-236	GEQVEFSFPLAFTVEKLTGSG	18549
F6 (45)	221-241	FSFPLAFTVEKLTGSGELWWQ	9673
G6 (46)	226-246	AFTVEKLTGSGELWWQAERAS	3992
H6 (47)	231-251	KLTGSGELWWQAERASSSKSW	1878
A7 (48)	236-256	GELWWQAERASSSKSWITFDL	2730
B7 (49)	241-261	QAERASSSKSWITFDLKNKEV	2588
C7 (50)	246-266	SSSKSWITFDLKNKEVSVKRV	1761
D7 (51)	251-271	WITFDLKNKEVSVKRVTQDPK	2126
E7 (52)	256-276	LKNKEVSVKRVTQDPKLQMKG	2288
F7 (53)	261-281	VSVKRVTQDPKLQMGGKKLPLH	1848
G7 (54)	266-286	VTQDPKLQMGGKKLPLHHTLPQ	2075
H7 (55)	271-291	KLQMGGKKLPLHHTLPQALPQY	1949
A8 (56)	276-296	KKLPLHHTLPQALPQYAGSGN	1922
B8 (57)	281-301	HLTLHQALPQYAGSGNLTLA	2394
C8 (58)	286-306	QALPQYAGSGNLTLAEEKTG	2364
D8 (59)	291-311	YAGSGNLTLAEEKTGKLHQE	1830
E8 (60)	296-316	NLTIAEEKTGKLHQEVNLVV	1676
F8 (61)	301-321	LEAKTGKLHQEVNLVVMRATQ	1729
G8 (62)	306-326	GKLHQEVNLVVMRATQLQKNL	1776
H8 (63)	311-331	EVNLVVMRATQLQKNLTCEVW	2183
A9 (64)	316-336	VMRATQLQKNLTCEVWGPTSP	2144
B9 (65)	321-341	QLQKNLTCEVWGPTSPKLMIL	1856
C9 (66)	326-346	LTCEVWGPTSPKLMILSLKLEN	2412
D9 (67)	331-351	WGPTSPKLMILSLKLENKEAKV	2414
E9 (68)	336-356	PKLMILSLKLENKEAKVSKREK	1656
F9 (69)	341-361	SLKLENKEAKVSKREKAVWVL	1663
G9 (70)	346-366	NKEAKVSKREKAVWVLNPEAG	1735
H9 (71)	351-371	VSKREKAVWVLNPEAGMWQCL	2034
A10 (72)	356-376	KAVWVLNPEAGMWQCLLSDSG	3133
B10 (73)	361-381	LNPEAGMWQCLLSDSGQVLLE	6316
C10 (74)	366-386	GMWQCLLSDSGQVLLESNIKV	4185
D10 (75)	371-391	LLSDSGQVLLESNIKVLPWTS	2375
E10 (76)	376-396	GQVLLESNIKVLPWSTPVQP	2089
F10 (77)	381-401	ESNIKVLPWSTPVQPMALIV	1992
G10 (78)	386-406	VLPTWSTPVQPMALIVLGGVA	2197
H10 (79)	391-411	STPVQPMALIVLGGVAGLLL	2527
A11 (80)	396-416	PMALIVLGGVAGLLLFIGLGI	3067
B11 (81)	401-421	VLGGVAGLLLFIGLGIFFCVR	3738
C11 (82)	406-426	AGLLLFIGLGIFFCVRCRHR	2099
D11 (83)	411-431	FIGLGIFFCVRCRHRQQAER	1900
E11 (84)	416-436	IFFCVRCRHRQQAERMSQIK	2085
F11 (85)	421-441	RCRHRRQQAERMSQIKRLLSE	2075
G11 (86)	426-446	RRQAERMSQIKRLLSEKKTCQ	1607

H11 (87)	431-451	RMSQIKRLLSEKKTQCQCPHRF	2020
A12 (88)	436-456	KRLLSEKKTCQCQCPHRFQKTCS	1674
B12 (89)	441-458	EKKTCQCQCPHRFQKTCSPI	2006
A1 ( 0 )		empty (control)	2075

### Example 5

This example shows the binding of  $^{125}\text{I}$ -HIV-1<sub>LAI</sub> gp120 to the amino termini of CCR5, CXCR4, and STRL33 as a function of the dependence on position and length. Synthetic peptide arrays of nonapeptides, dodecapeptides, pentadecapeptides and octadecapeptides derived from CCR5 (panel A), CXCR4 (panel B) and STRL33 (panel C) amino terminal domains were prepared and utilized to test the binding of  $^{125}\text{I}$ -HIV-1<sub>LAI</sub> envelope gp120. Ordinal sequence position numbers are given in accordance with the sequence data provided by the Genbank database for CCR5 (accession No. g1457946, gi|1457946), CXCR4 (accession No. g539677, gi|400654, sp|P30991) and STRL33 (accession No. g2209288, gi|2209288). The counts shown are the counts detected in each well minus the background counts (i.e., counts observed in the assay when no polypeptide was bound to the well of the 96-well assay plate).

Panel A Peptide Sequence Scanning Windows		Binding Results For Window Length			
CCRS	(In each sequence row 9-, Initial 12-, 15-, 18-mers share the Sequence same initial starting point.)	(counts bound - background (no peptide))			
#		9	9	12	15
	xxxxxxxxxx	9	9	12	15
	xxxxxxxxxxxxx	12			
	xxxxxxxxxxxxxxx	15			
	xxxxxxxxxxxxxxx	18			18
1	MDYQVSSPIYDINYYTSE	543	2682	4976	5880
2	DYQVSSPIYDINYYTSEP	1552	3089	5401	6363
3	YQVSSPIYDINYYTSEPC	2533	5305	5415	6119
4	QVSSPIYDINYYTSEPCQ	490	1959	4594	5645
5	VSSPIYDINYYTSEPCQK	509	1629	3280	3521
6	SSPIYDINYYTSEPCQKI	671	1739	3498	3285
7	SPIYDINYYTSEPCQKIN	1503	3463	4575	3234
8	PIYDINYYTSEPCQKINV	1186	2285	2682	2036
9	IYDINYYTSEPCQKINVK	1359	2702	2516	1261
10	YDINYYTSEPCQKINVKQ	4379	5245	3052	1913
11	DINYYTSEPCQKINVKQI	1396	1361	1144	712
12	INYYTSEPCQKINVKQIA	1384	1190	707	684
13	NYYTSEPCQKINVKQIAA	1548	977	760	595
14	YYTSEPCQKINVKQIAAR	1029	1052	847	638
15	YTSEPCQKINVKQIA	567	507	459	
16	TSEPCQKINVKQIAA	440	427	509	
17	SEPCQKINVKQIAAR	434	430	426	
18	EPCQKINVKQIA	397	432		
19	PCQKINVKQIAA	386	385		
20	CQKINVKQIAAR	435	581		
21	QKINVKQIA	453			
22	KINVKQIAA	487			
23	INVKQIAAR	474			

2020 RELEASE UNDER E.O. 14176

Panel B	Peptide Sequence Scanning Windows	Binding Results For Window Length			
CXCR4		(counts bound - background)			
Initial Sequence #	xxxxxx	9	9	12	15
	xxxxxxxx	12		12	
	xxxxxxxxxxxx	15			15
	xxxxxxxxxxxxxxxx	18			18
1	MEGISTYTSNDNYTEEMGS	591	334	3275	2079
2	EGISIYTSDNYTEEMGSG	a	886	7255	1548
3	GISIYTSDNYTEEMGSGD	454	2644	3274	1217
4	ISIYTSDNYTEEMGSGDY	466	3973	2202	861
5	SIYTSDNYTEEMGSGDYZ	a	288	168	239
6	IYTSNDNYTEEMGSGDYZDS	332	335	195	173
7	YTSNDNYTEEMGSGDYZDSM	181	161	201	103
8	TSDNYTEEMGSGDYZDSMK	a	54	119	38
9	SDNYTEEMGSGDYZDSMKE	151	149	124	161
10	DNYTEEMGSGDYZDSMKEP	67	121	57	102
11	NYTEEMGSGDYZDSMKEPC	a	100	30	134
12	YTEEMGSGDYZDSMKEPCF	68	213	70	103
13	TEEMGSGDYZDSMKEPCFR	146	67	23	47
14	EEMGSGDYZDSMKEPCFRE	a	61	121	130
15	EMGSGDYZDSMKEPCFREE	64	36	69	64
16	MGSGYDSMKEPCFREEEN	57	68	64	129
17	GSGDYDSMKEPCFREEENA	a	155	172	155
18	SGDYDSMKEPCFREEANAN	100	118	186	89
19	GDYDSMKEPCFREEANANF	53	167	198	134
20	DYDSMKEPCFREEANANFN	a	167	146	75
21	YDSMKEPCFREEANANFNK	171	144	80	89
22	DSMKEPCFREEANANFNKI	85	144	146	40
23	SMKEPCFREEANANFN	a	119	55	
24	MKEPCFREEANANFNK	188	133	74	
25	KEPCFREEANANFNKI	165	105	93	
26	EPCFREEANANFN	a	69		
27	PCFREEANANFNK	104	108		
28	CFREEANANFNKI	103	66		
29	REENANFNK	58			

a Not done

Panel C	Peptide Sequence Scanning Windows	Binding Results For Window Length			
STRL33		(counts bound - background)			
Initial Sequence #	(In each sequence row 9-, 12-, 15-, 18-mers share the same initial starting point.)				
	xxxxxx	9	9		
	xxxxxxxxxx	12		12	
	xxxxxxxxxxxx	15			15
	xxxxxxxxxxxxxx	18			18
1	MAEHDYHEDYGFSSFNDSS	160	625	1239	1386
2	AEHDYHEDYGFSSFNDSS	354	697	1095	1014
3	EHDYHEDYGFSSFNDSSQ	509	937	2235	1219
4	HDYHEDYGFSSFNDSSQE	708	1427	1772	1500
5	DYHEDYGFSSFNDSSQEE	851	1554	1240	1191
6	YHEDYGFSSFNDSSQEEH	728	1950	1357	985
7	HEDYGFSSFNDSSQEEHQ	729	1077	947	537
8	EDYGFSSFNDSSQEEHQAA	953	817	1152	548
9	DYGFSSFNDSSQEEHQAF	701	573	595	440
10	YGFSSFNDSSQEEHQAFQ	345	745	645	1138
11	GFFSFNDSSQEEHQAFQLQ	171	480	270	1639
12	FSSFNDSSQEEHQAFLQF	249	403	361	3608
13	SSFNDSSQEEHQAFLQFS	243	277	902	6038
14	SFNDSSQEEHQAFLQFSK	304	303	969	4537
15	FNDSSQEEHQAFLQFSKV	246	470	4089	4678
16	NDSSQEEHQAFLQFS	180	497	6160	
17	DSSQEEHQAFLQFSK	147	882	4588	
18	SSQEEHQAFLQFSKV	287	4455	4732	
19	SQEEHQAFLQFS	647	7512		
20	QEEHQAFLQFSK	1109	5672		
21	EEHQAFLQFSKV	6060	5598		
22	EHQAFLQFS	7505			
23	HQAFLQFSK	2761			
24	QAFLQFSKV	2600			

Example 6

This example shows  $^{125}\text{I}$ -HIV-1<sub>MA</sub> gp120 binding to  
5 N-terminal peptide variants of CCR5, CXCR4 and STRL33.

Octadecapeptide alanine replacement variants of maximum gp120 binding activity peaks were synthesized and tested for  $^{125}\text{I}$ -HIV-1<sub>LAI</sub> gp120 binding. Each binding value presented is the average of two separate synthesis and 5 binding experiments. Relative percentage of Control =  $\{[(\text{mean counts}/\text{Control counts})] \times 100\%\} \pm \text{average deviation}$ . Background counts (no peptide, see Example 7) were subtracted from all values. Data for CCR5 are presented in Panel A; data for CXCR4 are presented in 10 Panel B; and data for STRL33 are presented in Panel C.

Panel A.  $^{125}\text{I}$ -HIV-1<sub>LAI</sub> gp120 binding to N-terminal peptide variants of CCR5

	CCR5 variant peptides (1-18)	Relative % of Control <sup>a</sup>
Control	MDYQVSSPIYDINYYTSE	100
M1A	ADYQVSSPIYDINYYTSE	167 $\pm$ 4
D2A	MAYQVSSPIYDINYYTSE	125 $\pm$ 8
Y3A	MDAQVSSPIYDINYYTSE	51 $\pm$ 2
Q4A	MDYAVSSPIYDINYYTSE	104 $\pm$ 7
V5A	MDYQASSPIYDINYYTSE	82 $\pm$ 3
S6A	MDYQVASPIYDINYYTSE	124 $\pm$ 3
S7A	MDYQVSAPIYDINYYTSE	56 $\pm$ 2
P8A	MDYQVSSAIYDINYYTSE	157 $\pm$ 2
I9A	MDYQVSSPAYDINYYTSE	24 $\pm$ 7
Y10A	MDYQVSSPIADINYYTSE	19 $\pm$ 6
D11A	MDYQVSSPIYAINYYTSE	63 $\pm$ 22
I12A	MDYQVSSPIYDANYYTSE	14 $\pm$ 1
N13A	MDYQVSSPIYDIAYYTSE	253 $\pm$ 19
Y14A	MDYQVSSPIYDINAYTSE	15 $\pm$ 0.3
Y15A	MDYQVSSPIYDINYATSE	21 $\pm$ 5
T16A	MDYQVSSPIYDINYYASE	78 $\pm$ 34
S17A	MDYQVSSPIYDINYYTAE	64 $\pm$ 6
E18A	MDYQVSSPIYDINYYTSA	4 $\pm$ 2

<sup>a</sup>The percent binding for the wild-type peptide was defined as 100%.

Panel B       $^{125}\text{I}$ -HIV-1<sub>LAI</sub> gp120 binding to N-terminal peptide variants of  
CXCR4

	CXCR4 variant peptides (1-18)	Relative % of Control <sup>a</sup>
Control	MEGISIYTSNDNYTEEMGS	100
M1A	AEGISIYTSNDNYTEEMGS	118 ± 18
E2A	MAGISIYTSNDNYTEEMGS	36 ± 0.3
G3A	MEAISIYTSNDNYTEEMGS	101 ± 3
I4A	MEGASIYTSNDNYTEEMGS	6 ± 0.3
S5A	MEGIAIYTSNDNYTEEMGS	133 ± 5
I6A	MEGISAYTSNDNYTEEMGS	2 ± 1
Y7A	MEGISIATSDNYTEEMGS	7 ± 0.4
T8A	MEGISIYASDNYTEEMGS	97 ± 10
S9A	MEGISIYTADNYTEEMGS	70 ± 4
D10A	MEGISIYTSANYTEEMGS	71 ± 8
N11A	MEGISIYTSDAYTEEMGS	38 ± 0.4
Y12A	MEGISIYTSNDNATEEMGS	28 ± 2
T13A	MEGISIYTSNDNYAEEEMGS	70 ± 6
E14A	MEGISIYTSNDNYTAEMGS	72 ± 1
E15A	MEGISIYTSNDNYTEAMGS	56 ± 7
M16A	MEGISIYTSNDNYTEEAGS	88 ± 4
G17A	MEGISIYTSNDNYTEEMAS	68 ± 8
S18A	MEGISIYTSNDNYTEEMGA	79 ± 1

<sup>a</sup> The percent binding for the wild-type peptide was defined as 100%.

Panel C  $^{125}\text{I}$ -HIV-1<sub>LAI</sub> gp120 binding to N-terminal peptide variants of  
STRL33

	STRL33 variant peptides (21-38)	Relative % of Control <sup>a</sup>
Control	EEHQ AFLQFSK VFLPCM Y	100
E21A	AEHQ AFLQFSK VFLPCM Y	81 ± 2
E22A	EAHQ AFLQFSK VFLPCM Y	70 ± 1
H23A	EEAQ AFLQFSK VFLPCM Y	99 ± 1
Q24A	EEHA AFLQFSK VFLPCM Y	72 ± 1
A25A	EEHQ AFLQFSK VFLPCM Y	101 ± 1
F26A	EEHQ AALQFSK VFLPCM Y	32 ± 0.1
L27A	EEHQ AF AQFSK VFLPCM Y	37 ± 2
Q28A	EEHQ AFLAFSK VFLPCM Y	44 ± 0.4
F29A	EEHQ AFLQASK VFLPCM Y	20 ± 1
S30A	EEHQ AFLQFAK VFLPCM Y	92 ± 2
K31A	EEHQ AFLQFS A VFLPCM Y	162 ± 2
V32A	EEHQ AFLQFSKA F LPCM Y	51 ± 3
F33A	EEHQ AFLQFSK V ALPCM Y	45 ± 2
L34A	EEHQ AFLQFSK V FAPCM Y	76 ± 1
P35A	EEHQ AFLQFSK V FLACM Y	82 ± 3
C36A	EEHQ AFLQFSK V FLPAMY	53 ± 5
M37A	EEHQ AFLQFSK V FLCAY	112 ± 4
Y38A	EEHQ AFLQFSK V FLCMA	83 ± 2

<sup>a</sup> The percent binding for the wild-type peptide was defined as 100%.

Example 7

- 5 This example demonstrates that the binding of HIV-1 gp120 envelope protein to the polypeptides of the present invention and to the chemokine receptors from which the present inventive polypeptides were originally derived or inspired is conserved across the various species of
- 10 HIV-1. This example also demonstrates that a step subsequent to initial binding of gp120 to CCR5, CXCR4, STRL33, and CD4 is the most likely source of the phenomenon of host-range selectivity. Additionally, this example demonstrates that the underlying method is
- 15 accurate in that receptor variants that are predicted to have an altered affinity for binding with gp120, do in

fact have a statistically similar alteration in affinity where comparable changes in the receptors have been identified in other work and the affinity for binding of gp120/effect on infectivity has been measured.

5 This example examines the effect of particular mutations of CCR5 that were studied in the work underlying the present invention and that were also studied by other artisans in the field.

The following table identifies a mutation in the  
10 first column. The first letter designates the wild-type amino acid present at the position indicated by the number, and the letter A which terminates all entries in the first column indicates that the amino acid residue present in that position in the mutant polypeptide is  
15 alaninyl. For example, the first data row (i.e., the second row of the table) contains the entry Y3A in the first column, which indicates that the tyrosine residue at position 3 of the wild-type CCR5 is substituted by an alanine residue.

20 The second column provides the percentage of binding exhibited by a mutant polypeptide compared to a wild-type polypeptide, when the methods used to elucidate the present invention are used in conjunction with radiolabeled HIV-1<sub>LA1</sub> gp120 envelope protein. The third through seventh columns provide similar data that have been extracted from the work of others in the field using a strain of HIV-1 virus indicated at the top of each column. For example, row 2 of the following table indicates that when the mutation Y3A is effected in the  
25 human CCR5 chemokine receptor, then the resulting CCR5 polypeptide has 51.4% of the ability to bind HIV-1<sub>LA1</sub>

gp120 envelope protein in comparison to an equivalent wild-type peptide. Similarly, HIV-1<sub>ADA</sub> binds to the mutant polypeptide with 79% of the affinity of a non-mutated CCR5 chemokine receptor.

5

	gp120	YU2	ADA	JF-RL	89.6	DH123
Y3A	51.4	n/a	79	82	n/a	42
Q4A	104	85	132	111	67	105
Y10A	19.2	2	50	26	10	3
D11A	62.8	2	27	22	6	3
Y14A	14.6	12	47	25	6	0
Y15A	21	30	3	3	1	0
E18A	4.1	45	12	12	3	10

Statistical analysis of these data indicates that the similarity between the binding affinity of each mutant peptide for gp120 elucidated in this study is not more than about 25% likely to be causally unrelated to the effects observed for YU2, and not more than about 4% likely to be causally unrelated to the effects observed for each of the other viruses listed in the table above.

Additionally, the affinity measurements generated by the underlying technique has been demonstrated to be accurate by (repetitively) showing that antibodies that specifically bind to radiolabeled gp120 are capable of preventing the binding of gp120 to polypeptides that have shown high affinity for binding with gp120 in the experiments upon which the present invention is predicated. Thus, this example shows that the binding with chemokine receptors HIV-1 can be inhibited by the present inventive polypeptides, irrespective of the strain of HIV-1 from which the gp120 protein is obtained.

Example 8

This example provides a characterization of the critical amino acids in the amino-terminal segments of CCR5, CXCR4, and STRL33 that are essential for the ability of these polypeptides to bind with gp120.

In this example, the effect on binding that occurs to due successive replacement of each amino acid with alanine is indicated, wherein a (+) signifies a decrease in binding affinity and a (>) signifies an enhancement in binding affinity. As is clear from inspection, the sequences are shown with that amino-terminus at top and the carboxyl-terminus at bottom.

<u>CCR5 (1-18)</u>	<u>CXCR4 (1-18)</u>	<u>STRL33 (21-38)</u>
M>	M	E
D	E+	E
Y++	G	H
Q	I+++++	Q
V	S>	A
S	I++++++	F+++
S+	Y+++++	L++
P>	T	Q+
I+++	S+	F+++
Y+++	D+	S
D+	N++	K>
I++++	Y++	V+
N>	T	F+
Y++++	E	L
Y+++	E++	P
T	M	C+
S+	G	M
E+++++	S	Y

Example 9

This example employs the same technique as Example 4 and provides information similar to that available from Example 4.

5       The data below compares the ability of synthetic fragments of CD4 to bind to labeled gp120. 9-mer, 12-mer, 15-mer, 18-mer, and 21-mers were selected based on the data from Examples 4. The relative binding affinities of each group of polypeptides can be 10 determined by inspection of the number of counts of radiolabeled gp120 that were retained by each N-mer. Data supporting these conclusions are provided by Examples 10 and 11.

Peptide starting position #	Active Peptides	gp120 bound (counts)	Peptide starting position #	Active Peptides	Gp120 Bound (counts)
	<u>ACTIVE 9-MERS</u>			<u>ACTIVE 12-MERS</u>	
105	DTYICEVED	1043	101	IEDSDTYICEVE	1107
115	KEEVQLLVF	1273	112	EDQKEEVQLLVF	1379
116	EEVQLLVFG	3170	113	DQKEEVQLLVFG	1624
117	EVQLLVFGL	2146	114	QKEEVQLLVFGL	1785
			115	KEEVQLLVFGLT	1774
			116	EEVQLLVFGLTA	3261
			117	EVQLLVFGLTAN	1838
			133	LLQGQSLTLTLE	1320
217	EQVEFSFPL	1032	215	EGEQVEFSFPLA	1456
218	QVEFSFPLA	1205	216	GEQVEFSFPLAF	1729
219	VEFSFPLAF	1064	217	EQVEFSFPLAFT	1556
			218	QVEFSFPLAFTV	1636
	<u>ACTIVE 15-MERS</u>			<u>ACTIVE 18-MERS</u>	
109	CEVEDQKEEVQLLVF	1729	105	DTYICEVEDQKEE	1648
110	EVEDQKEEVQLLVFG	2805	106	VQLLV TYICEVEDQKEEV	3794
111	VEDQKEEVQLLVFGL	3816	107	QLLVF YICEVEDQKEEVQ	4611

			LLVFG	
112	EDQKEEVQLLVFGLT	3633	108 ICEVEDQKEEVQL	3898
113	DQKEEVQLLVFGLTA	3905	109 LVFGL	
114	QKEEVQLLVFGLTAN	3770	109 CEVEDQKEEVQLL	3797
115	KEEVQLLVFGLTANS	3485	110 VFGLT	
116	EEVQLLVFGLTANS	6423	110 EVEDQKEEVQLLV	3647
117	EVQLLVFGLTANS	2689	111 FGLTA	
117	EVQLLVFGLTANS	2689	111 VEDQKEEVQLLVF	3913
117	EVQLLVFGLTANS	2689	112 GLTAN	
117	EVQLLVFGLTANS	2689	112 EDQKEEVQLLVFG	3416
117	EVQLLVFGLTANS	2689	112 LTANS	
117	EVQLLVFGLTANS	2689	113 DQKEEVQLLVFGL	3317
117	EVQLLVFGLTANS	2689	113 TANS	
117	EVQLLVFGLTANS	2689	114 QKEEVQLLVFGLT	3671
117	EVQLLVFGLTANS	2689	114 ANSDT	
130	DTHLLQGQSLTLTLE	1622	127 ANSDTHLLQGQSL	1540
131	THLLQGQSLTLTLES	1874	127 TLTLE	
132	HLLQGQSLTLTLES	1277	128 NSDTHLLQGQSLT	1726
132	HLLQGQSLTLTLES	1277	128 LTLES	
213	KKEGEQVEFSFPLAF	1921	129 SDTHLLQGQSLTL	1260
213	KKEGEQVEFSFPLAF	1921	129 TLESP	
214	KEGEQVEFSFPLAFT	3253	210 IVYKKEGEQVEFS	5382
214	KEGEQVEFSFPLAFT	3253	210 FPLAF	
215	EGEQVEFSFPLAFTV	3270	211 VYKKEGEQVEFSF	4307
215	EGEQVEFSFPLAFTV	3270	211 PLAFT	
216	GEOQVEFSFPLAFTVE	4656	212 YKKEGEQVEFSFP	4839
216	GEOQVEFSFPLAFTVE	4656	212 LAFTV	
217	EQVEFSFPLAFTVEK	4135	213 KKEGEQVEFSFPL	4683
217	EQVEFSFPLAFTVEK	4135	213 AFTVE	
218	QVEFSFPLAFTVEKL	2047	214 KEGEQVEFSFPLA	3117
218	QVEFSFPLAFTVEKL	2047	214 FTVEK	
			215 EGEQVEFSFPLAF	2164
			215 TVEKL	
			216 GEOQVEFSFPLAFT	1643
			216 VEKLT	
			ACTIVE 21-MERS	
90	GNFPLIIKNLKIEDS	5248		
	DTYICE			
91	NFPLIIKNLKIEDSD	7803		
	TYICEV			

2022-07-20 14:53:00

92	FPLIIKNLKIEDSDT YICEVE	13919
93	PLIIKNLKIEDSDTY ICEVED	20145
94	LIIKNLKIEDSDTYI CEVEDQ	17108
95	IICKNLKIEDSDTYIC EVEDQK	11892
96	IKNLKIEDSDTYICE VEDQKE	15073
97	KNLKIEDSDTYICEV EDQKEE	8789
99	LKIEDSDTYICEVED QKEEVQ	5519
100	KIEDSDTYICEVEDQ KEEVQL	6325
101	IEDSDTYICEVEDQK EEVQLL	12064
102	EDSDTYICEVEDQKE EVQLLV	4933
103	DSDTYICEVEDQKEE VQLLVF	30277
104	SDTYICEVEDQKEEV QLLVFG	30319
105	DTYICEVEDQKEEVQ LLVFGGL	25424
106	TYICEVEDQKEEVQL LVFGLT	20191
107	YICEVEDQKEEVQLL VFGLTA	22884
108	ICEVEDQKEEVQLLV FGLTAN	7276
109	CEVEDQKEEVQLLVF GLTANS	3517
123	FGLTANS DTHLL QGQ SLTL TL	11529
124	GLTANS DTHLL QGQS LT LTLE	14065
125	LTANS DTHLL QGQSL TL TLES	17113
126	TANS DTHLL QGQSLT LT LESP	23595

204	FQKASSIVYKKEGEQ VEFSFP	9382
205	QKASSIVYKKEGEQV EFSFPL	24959
206	KASSIVYKKEGEQEVE FSFPLA	30873
207	ASSIVYKKEGEQEVEF SFPLAF	25146
208	SSIVYKKEGEQEVEFS FPLAFT	28068
209	SIVYKKEGEQEVEFSF PLAFTV	8165
210	IVYKKEGEQEVEFSFP LAFTVE	15620
221	FSFPLAFTVEKLTGS GELWWQ	4163
222	SFPLAFTVEKLTGSG ELWWQA	2284
223	FPLAFTVEKLTGSGE LWWQAE	6276
224	PLAFTVEKLTGSGEL WWQAER	2647
225	LAFTVEKLTGSGELW WQAERA	3577

Example 10.

This example provides data which enables those skilled in the art to arrive at the conclusions indicated in Examples 9 and 12. In this example, the counts of radiolabeled gp-120 retained by each peptide indicated in the left hand column are given in the right hand column. The first panel (panel A) provides data for 21-mers of CD4.

10

Panel A  
PEPTIDE

COUNTS

LWDQGNFPLIIKNLKIEDSDT	731
WDQGNFPLIIKNLKIEDSDTY	889
DQGNFPLIIKNLKIEDSDTYI	1138

QGNFPLIIKNLKIEDSDTYIC	2242
GNFPLIIKNLKIEDSDTYICE	5248
NFPLIIKNLKIEDSDTYICEV	7803
FPLIIKNLKIEDSDTYICEVE	13919
PLIIKNLKIEDSDTYICEVED	20145
LIIKNLKIEDSDTYICEVEDQ	17108
IICKNLKIEDSDTYICEVEDQK	11892
IKNLKIEDSDTYICEVEDQKE	15073
KNLKIEDSDTYICEVEDQKEE	8789
NLKIEDSDTYICEVEDQKEEV	2016
LKIEDSDTYICEVEDQKEEVQ	5519
IEDSDTYICEVEDQKEEVQL	6325
EDSDTYICEVEDQKEEVQLLV	12064
DSDTYICEVEDQKEEVQLLVF	4933
SDTYICEVEDQKEEVQLLVFG	30277
DTYICEVEDQKEEVQLLVGLT	30319
YICEVEDQKEEVQLLVGLTA	25424
ICEVEDQKEEVQLLVGLTAN	20191
CEVEDQKEEVQLLVGLTANS	22884
EVEDQKEEVQLLVGLTANS	7276
VEDQKEEVQLLVGLTANS	3517
EDQKEEVQLLVGLTANS	1687
DQKEEVQLLVGLTANS	646
QKEEVQLLVGLTANS	562
KEEVQLLVGLTANS	599
EEVQLLVGLTANS	573
EVQLLVGLTANS	682
VQLLVGLTANS	690
QLLVFGLTANS	589
LLVFGLTANS	1099
LVFGLTANS	2057
VFGLTANS	860
FGLTANS	4677
GLTANS	2762
LTANS	11529
TANS	14065
Empty (Control)	17113
TWTCTVLQNQKKVEFKIDIVV	23595
WTCTVLQNQKKVEFKIDIVVL	515
TCTVLQNQKKVEFKIDIVVLA	1430
CTVLQNQKKVEFKIDIVVLA	1616
TVLQNQKKVEFKIDIVVLA	1092
VLQNQKKVEFKIDIVVLA	2909
QK	3273
	1323

LQNQKKVEFKIDIVVLA	1256
QNQKKVEFKIDIVVLA	1808
NQKKVEFKIDIVVLA	1507
QKKVEFKIDIVVLA	759
KKVEFKIDIVVLA	782
KVEFKIDIVVLA	635
VEFKIDIVVLA	725
EFKIDIVVLA	649
FKIDIVVLA	593
KIDIVVLA	1394
IDIVVLA	962
DIVVLA	788
IVVLA	646
VVLAFQKASSIVY	772
VLA	1793
LAFQKASSIVY	1410
AFQKASSIVY	3775
FQKASSIVY	9382
QKASSIVY	24959
KASSIVY	30873
ASSIVY	25146
SSIVY	28068
SIVY	8165
IVY	15620
VYK	2429
YK	735
KKEGEQVEFS	1847
KEGEQVEFS	972
EGEQVEFS	739
GEQVEFS	652
EQVEFS	765
QVEFS	741
VEFS	633
EFS	681
FSF	4163
SF	2284
FPL	6276
PLA	2647
LAFT	3577
AFT	1739
Empty (control)	617

These second and third panels (panels B and C) provide data for 18-mers of a small region of CD4.

## Panel B

## PEPTIDE

## COUNTS

LWDQGNFPLIIKNLK	502
WDQGNFPLIIKNLKI	534
DQGNFPLIIKNLKIE	635
QGNFPLIIKNLKIED	509
GNFPLIIKNLKIEDS	624
NFPLIIKNLKIEDSD	654
FPLIIKNLKIEDSDT	539
PLIICKNLKIEDSDTY	661
LIIKNLKIEDSDTYI	542
IICKNLKIEDSDTYIC	664
IKNLKIEDSDTYICE	568
KNLKIEDSDTYICEV	562
NLKIEDSDTYICEVE	1160
LKIEDSDTYICEVED	846
KIEDSDTYICEVEDQ	1088
IEDSDTYICEVEDQK	1143
EDSDTYICEVEDQKE	815
DSDTYICEVEDQKEE	973
SDTYICEVEDQKEEV	993
DTYICEVEDQKEEVQ	1071
TYICEVEDQKEEVQL	956
YICEVEDQKEEVQLL	1064
ICEVEDQKEEVQLLV	1084
CEVEDQKEEVQLLVF	1729
EVEDQKEEVQLLVFG	2805
VEDQKEEVQLLVFGL	3816
EDQKEEVQLLVFGLT	3633
DQKEEVQLLVFGLTA	3905
QKEEVQLLVFGLTAN	3770
KEEVQLLVFGLTANS	3485
EEVQLLVFGLTANS	6423
EVQLLVFGLTANS	2689
VQLLVFGLTANS	1006
QLLVFGLTANS	865
LLVFGLTANS	599
LVFGLTANS	609
VFGLTANS	532
FGLTANS	625

GLTANSDTLLLQGQS	532
LTANSDTLLLQGQSL	634
TANSDTLLLQGQSLT	513
ANSDTLLLQGQSLTL	542
NSDTLLLQGQSLTLT	631
SDTLLLQGQSLTLTL	747
DTHLLQGQSLTLTLE	1622
THLLQGQSLTLTLES	1874
HLLQGQSLTLTLESP	1277
LWDQGNFPLIIKNLKIED	582
WDQGNFPLIIKNLKIEDS	626
DQGNFPLIIKNLKIEDSD	598
QGNFPLIIKNLKIEDSDT	564
GNFPLIIKNLKIEDSDTY	557
NFPLIIKNLKIEDSDTYI	627
FPLIIKNLKIEDSDTYIC	509
PLIIKNLKIEDSDTYICE	624
LIIKNLKIEDSDTYICEV	634
IICKNLKIEDSDTYICEVE	751
IKNLKIEDSDTYICEVED	699
KNLKIEDSDTYICEVEDQ	708
NLKIEDSDTYICEVEDQK	863
LKIEDSDTYICEVEDQKE	872
KIEDSDTYICEVEDQKEE	858
IEDSDTYICEVEDQKEEV	1230
EDSDTYICEVEDQKEEVQ	788
DSDTYICEVEDQKEEVQL	961
SDTYICEVEDQKEEVQLL	870
DTYICEVEDQKEEVQLLV	1648
TYICEVEDQKEEVQLLVF	3794
YICEVEDQKEEVQLLVFG	4611
ICEVEDQKEEVQLLVFGL	3898
CEVEDQKEEVQLLVFGLT	3797
EVEDQKEEVQLLVFGLTA	3647
VEDQKEEVQLLVFGLTAN	3913
EDQKEEVQLLVFGLTANS	3416
DQKEEVQLLVFGLTANS	3317
QKEEVQLLVFGLTANS	3671
KEEVQLLVFGLTANS	1271
EEVQLLVFGLTANS	783
EVQLLVFGLTANS	667
VQLLVFGLTANS	673
QLLVFGLTANS	574
LLVFGLTANS	568
LVFGLTANS	564
LVFGLTANS	564

65

VFGLTANS DTHLLQGQSL	531
FGLTANS DTHLLQGQSLT	591
GLTANS DTHLLQGQSLTL	572
LTANS DTHLLQGQSLTLT	528
TANS DTHLLQGQSLTLT	891
ANS DTHLLQGQSLTLTLE	1540
NSDTHLLQGQSLTLTLES	1726
SDTHLLQGQSLTLTLESP	1260
Empty (control)	575

## Panel C

PEPTIDE	COUNTS
WTCTVLQNQKKVEFK	566
TCTVLQNQKKVEFKI	510
CTVLQNQKKVEFKID	608
TVLQNQKKVEFKIDI	587
VLQNQKKVEFKIDIV	605
LQNQKKVEFKIDIVV	644
QNQKKVEFKIDIVVL	636
NQKKVEFKIDIVVLA	860
QKKVEFKIDIVVLA F	1333
KKVEFKIDIVVLA FQ	951
KVEFKIDIVVLA FQK	1051
VEFKIDIVVLA FQKA	1005
EFKIDIVVLA FQKAS	1188
FKIDIVVLA FQKASS	1001
KIDIVVLA FQKASSI	956
IDIVVLA FQKASSIV	865
DIVVLA FQKASSIVY	776
IVVLA FQKASSIVYK	783
VVLA FQKASSIVYKK	577
VLA FQKASSIVYKKE	634
LAFQKASSIVYKKEG	593
AFQKASSIVYKKEGE	544
FQKASSIVYKKEGEQ	637
QKASSIVYKKEGEQV	519
KASSIVYKKEGEQVE	563
ASSIVYKKEGEQVEF	589
SSIVYKKEGEQVEF S	558
SIVYKKEGEQVEFSF	651
IVYKKEGEQVEFSFP	615
VYKKEGEQVEFSFPL	714

YKKEGEQVEFSFPLA	687
KKEGEQVEFSFPLAF	1921
KEGEQVEFSFPLAFT	3253
EGEQVEFSFPLAFTV	3270
GEQVEFSFPLAFTVE	4656
EQVEFSFPLAFTVEK	4135
QVEFSFPLAFTVEKL	2047
VEFSFPLAFTVEKLT	899
EFSFPLAFTVEKLTG	920
FSFPLAFTVEKLTGS	672
SFPLAFTVEKLTGSG	565
FPLAFTVEKLTGSGE	556
PLAFTVEKLTGSGEL	612
LAFTVEKLTGSGELW	579
AFTVEKLTGSGELWW	586
FTVEKLTGSGELWWQ	625
TVEKLTGSGELWWQA	550
VEKLTGSGELWWQAE	735
EKLTGSGELWWQAER	683
WTCTVLQNQKKVEFKIDI	588
TCTVLQNQKKVEFKIDIV	571
CTVLQNQKKVEFKIDIVV	553
TVLQNQKKVEFKIDIVVL	655
VLQNQKKVEFKIDIVVLA	724
LQNQKKVEFKIDIVVLAF	938
QNQKKVEFKIDIVVLAFQ	917
NQKKVEFKIDIVVLAFQK	889
QKKVEFKIDIVVLAFQKA	1013
KKVEFKIDIVVLAFQKAS	912
KVEFKIDIVVLAFQKASS	1011
VEFKIDIVVLAFQKASSI	819
EFKIDIVVLAFQKASSIV	799
FKIDIVVLAFQKASSIVY	843
KIDIVVLAFQKASSIVYK	779
IDIVVLAFQKASSIVYKK	711
DIVVLAFQKASSIVYKKE	660
IVVLAFQKASSIVYKKEG	531
VVLAFQKASSIVYKKEGE	560
VLAFLQKASSIVYKKEGEQ	549
LAFQKASSIVYKKEGEQV	665
AFQKASSIVYKKEGEQE	514
FQKASSIVYKKEGEQVEF	528
QKASSIVYKKEGEQVEFS	602
KASSIVYKKEGEQVEFSF	536
ASSIVYKKEGEQVEFSFP	701

67

SSIVYKKEGEQVEFSFPL	756
SIVYKKEGEQVEFSFPLA	771
IVYKKEGEQVEFSFPLAF	5382
VYKKEGEQVEFSFPLAFT	4307
YKKEGEQVEFSFPLAFTV	4839
KKEGEQVEFSFPLAFTVE	4683
KEGEQVEFSFPLAFTVEK	3117
EGEQVEFSFPLAFTVEKL	2164
GEQVEFSFPLAFTVEKLT	1643
EQVEFSFPLAFTVEKLTG	798
QVEFSFPLAFTVEKLTGS	736
VEFSFPLAFTVEKLTGSG	533
EFSFPLAFTVEKLTGSGE	668
FSFPLAFTVEKLTGSGEL	613
SFPLAFTVEKLTGSGELW	656
FPLAFTVEKLTGSGELWW	586
PLAFTVEKLTGSGELWWQ	650
LAFTVEKLTGSGELWWQA	866
AFTVEKLTGSGELWWQAЕ	788
FTVEKLTGSGELWWQAЕR	1143
Empty (control)	556

The fourth and fifth panels (Panels D and E) provide data for select 9-mers and 12-mers of CD4.

**5 Panel D PEPTIDE COUNTS**

DQGNFPLII	662
QGNFPLIIK	508
GNFPLIIKN	600
NFPLIIKNL	561
FPLIIKNLK	601
PLIINKLKI	697
LIIKNLKIE	515
IICKNLKIED	658
IKNLKIEDS	557
KNLKIEDSD	612
NLKIEDSDT	512
LKIEDSDTY	492
KIEDSDTYI	603
IEDSDTYIC	567
EDSDTYICE	650
DSDTYICEV	712

SDTYICEVE	819
DTYICEVED	1043
TYICEVEDQ	805
YICEVEDQK	728
ICEVEDQKE	596
CEVEDQKEE	555
EVEDQKEEV	587
VEDQKEEVQ	521
EDQKEEVQL	564
DQKEEVQLL	589
QKEEVQLLV	636
KEEVQLLVF	1273
EEVQLLVFG	3170
EVQLLVFGL	2146
VQLLVFGLT	815
QLLVFGLTA	822
LLVFGLTAN	576
LVFGLTANS	522
VFGLTANS	549
FGLTANS	563
GLTANS	481
LTANS	596
TANS	554
ANS	642
NSDTHLLQ	561
NSDTHLLQG	526
SDTHLLQGQ	578
DTHLLQGQS	512
THLLQGQSL	564
HLLQGQSLT	568
LLQGQSLTL	501
LQGQSLTLT	594
QGQSLTLTL	777
GQSLTLTLE	604
DQGNFPLIIKNL	533
QGNFPLIIKNLK	547
GNFPLIIKNLKI	647
NFPLIIKNLKIE	511
FPLIIKNLKIED	565
PLIIKNLKIEDS	619
LIIKNLKIEDSD	511
IICKNLKIEDSDT	574
IKNLKIEDSDTYI	523
KNLKIEDSDTYIC	639
NLKIEDSDTYIC	635
LKIEDSDTYICE	

KIEDSDTYICEV	601
IEDSDTYICEVE	1107
EDSDTYICEVED	956
DSDTYICEVEDQ	937
SDTYICEVEDQK	846
DTYICEVEDQKE	720
TYICEVEDQKEE	818
YICEVEDQKEEV	734
ICEVEDQKEEVQ	585
CEVEDQKEEVQL	561
EVEDQKEEVQLL	508
VEDQKEEVQLLV	657
EDQKEEVQLLVF	1379
DQKEEVQLLVFG	1624
QKEEVQLLVFGL	1785
KEEVQLLVFGLT	1774
EEVQLLVFGLTA	3261
EVQLLVFGLTAN	1838
VQLLVFGLTANS	747
QLLVFGLTANS	721
LLVFGLTANS	533
LVFGLTANS	586
VFGLTANS	548
FGLTANS	571
GLTANS	574
LTANS	534
TANS	549
ANS	559
NSDTHLLQGQS	585
NSDTHLLQGQSL	540
SDTHLLQGQSLT	527
DTHLLQGQSLTL	646
THLLQGQSLTL	701
HLLQGQSLTL	1320
LLQGQSLTL	581
Empty (control)	

## Panel E

PEPTIDE COUNTS

TVLQNQKKV	534
VLQNQKKVE	556
LQNQKKVEF	565
QNQKKVEFK	537
NQKKVEFKI	597

40084619 322702

QKKVEFKID	575
KKVEFKIDI	501
KVEFKIDIV	555
VEFKIDIVV	548
EFKIDIVVL	665
FKIDIVVLA	568
KIDIVVLAFL	665
IDIVVLAFLQ	691
DIVVLAFLQK	686
IVVLAFLQKA	602
VVLAFLQKAS	600
VLAFLQKASS	466
LAFQKASSI	592
AFQKASSIV	595
FQKASSIVY	568
QKASSIVYK	494
KASSIVYKK	498
ASSIVYKKE	600
SSIVYKKEG	515
SIVYKKEGE	566
IVYKKEGEQ	534
VYKKEGEQV	490
YKKEGEQVE	518
KKEGEQVEF	546
KEGEQVEFS	595
EGEQVEFSF	735
GEQVEFSFP	697
EQVEFSFPL	1032
QVEFSFPLA	1205
VEFSFPLAF	1064
EFSFPLAFT	658
FSFPLAFTV	472
SFPLAFTVE	619
FPLAFTVEK	569
PLAFTVEKL	597
LAFTVEKLT	501
AFTVEKLTG	517
FTVEKLTGS	574
TVEKLTGSG	487
VEKLTGSGE	585
EKLTGSSEL	541
KLTGSSELW	491
LTGSSELWW	550
TGSSELWWQ	507
TVLQNQKKVEFK	563

VLQNQKKVEFKI	503
LQNQKKVEFKID	508
QNQKKVEFKIDI	559
NQKKVEFKIDIV	532
QKKVEFKIDIVV	595
KKVEFKIDIVVL	597
KVEFKIDIVVLA	560
VEFKIDIVVLA	681
EFKIDIVVLA	659
FKIDIVVLA	736
KDIVVLA	689
IDIVVLA	630
DIVVLA	746
IVVLA	548
VVLAFQKASSIV	567
VLA	548
FQKASSIVY	465
AFQKASSIVY	597
FQKASSIVYK	577
QKASSIVYK	596
KASSIVYK	559
ASSIVYK	523
SSIVYK	615
SIVYK	543
IVYK	533
VYK	584
YK	548
KKEGEQVEFS	598
KKEGEQVEFSF	710
KEGEQVEFSFP	1456
EQEVS	1729
EQVEFS	1556
QVEFS	1636
VEFS	518
EFS	585
FSF	573
SF	528
FPL	622
PLAFT	528
VEKL	608
TGSG	511
AFT	530
FTV	573
TVE	477
VEKLTGSGELWW	543
EKLTGSGELWWQ	

Empty 571  
(control)

Panels F and G provide data on sequential alanine replacements for selected CD4 polypeptides.

5 Panel F

PEPTIDE	COUNTS
ZZZZZZZDTYICEVED	5844
ZZZZZZZATYICEVED	5921
ZZZZZZZDAYICEVED	6362
ZZZZZZZDTAICEVED	1301
ZZZZZZZDTYACEVED	2583
ZZZZZZZDTYIAEVED	4483
ZZZZZZZDTYICAVEAD	3154
ZZZZZZZDTYICEAED	3432
ZZZZZZZDTYICEVAD	3595
ZZZZZZZDTYICEVEA	5942
ZZZZZZZDTYICEVED	4973
ZZZZZZZDTYICEVED	4775
ZZZZZZZATYICEVED	4962
ZZZZZZZDAYICEVED	4163
ZZZZZZZDTAICEVED	1384
ZZZZZZZDTYACEVED	3085
ZZZZZZZDTYIAEVED	5128
ZZZZZZZDTYICAVEAD	2587
ZZZZZZZDTYICEAED	2499
ZZZZZZZDTYICEVAD	2706
ZZZZZZZDTYICEVEA	6345
ZZZZZZZDTYICEVED	5564
EEVQLLVFGLTANS	18582
AEVQLLVFGLTANS	16220
EAVQLLVFGLTANS	14220
EEAQQLVFGLTANS	18124
EEVALLVFGLTANS	10890
EEVQALVFGLTANS	11258
EEVQLAVFGLTANS	11954
EEVQLLAFGLTANS	13317
EEVQLLVAGLTANS	9573
EEVQLLVFALTANS	19348
EEVQLLVFGATANS	10408
EEVQLLVFGLAANS	19973

EEVQLLVFGLTNSD	20100
EEVQLLVFGLTAASD	19390
EEVQLLVFGLTANAD	17684
EEVQLLVFGLTANSA	18227
EEVQLLVFGLTANS'D	19738
EEVQLLVFGLTANS'D	21338
AEVQLLVFGLTANS'D	14590
EAVQLLVFGLTANS'D	13213
EEAQLLVFGLTANS'D	16296
EEVALLVFGLTANS'D	13415
EEVQALVFGLTANS'D	12603
EEVQLAVFGLTANS'D	13690
EEVQLLAFGLTANS'D	16286
EEVQLLVAGLTANS'D	11480
EEVQLLVFALTANS'D	18254
EEVQLLVFGATANS'D	19978
EEVQLLVFGLAANS'D	18863
EEVQLLVFGLTNSD	20021
EEVQLLVFGLTAASD	19200
EEVQLLVFGLTANAD	17928
EEVQLLVFGLTANSA	22206
EEVQLLVFGLTANS'D	18721
THLLQGQSLTLTLES	7756
AHLLQGQSLTLTLES	8602
TALLQGQSLTLTLES	6931
THALQGQSLTLTLES	7683
THLAQGQSLTLTLES	7701
THLLAGQSLTLTLES	4578
THLLQAQSLTLTLES	8471
THLLQGASLTTLTLES	4238
THLLQGQALTLTLES	8659
THLLQGQSATLTLES	4430
THLLQGQSLALTLES	8158
THLLQGQSLTATLES	4380
THLLQGQSLTLALES	11699
THLLQGQSLTLTAE'S	862
THLLQGQSLTLTLEAS	2596
THLLQGQSLTLTLEA	5849
THLLQGQSLTLTLES	6545
THLLQGQSLTLTLES	4787
AHLLQGQSLTLTLES	5826
TALLQGQSLTLTLES	5012
THALQGQSLTLTLES	5059
THLAQGQSLTLTLES	5120
THLLAGQSLTLTLES	2956

THLLQAQSLTLTLES	6393
THLLQGASLTTLTLES	1933
THLLQGQALTTLTLES	5151
THLLQGQSATTLTLES	1391
THLLQGQSLALTLES	4749
THLLQGQSLTATLES	813
THLLQGQSLTLALES	8147
THLLQGQSLTTLAES	797
THLLQGQSLTTLAS	2193
THLLQGQSLTTLTEA	7984
THLLQGQSLTTLTES	5947
Empty (control)	569

Panel G

PEPTIDE	COUNTS
GEQVEFSFPLAFTVE	20691
AEQVEFSFPLAFTVE	18546
GAQVEFSFPLAFTVE	17733
GEAVEFSFPLAFTVE	17500
GEQAEFSFPLAFTVE	14764
GEQVAFSFPLAFTVE	16668
GEQVEASFPLAFTVE	6793
GEQVEFAFPLAFTVE	21681
GEQVEFSAPLAFTVE	7767
GEQVEFSFALRAFTVE	20480
GEQVEFSFPAAFTVE	10024
GEQVEFSFPLTFTVE	17397
GEQVEFSFPLAATVE	10130
GEQVEFSFPLAFAVE	20627
GEQVEFSFPLAFTAE	18797
GEQVEFSFPLAFTVA	18371
GEQVEFSFPLAFTVE	17662
GEQVEFSFPLAFTVE	19190
AEQVEFSFPLAFTVE	18042
GAQVEFSFPLAFTVE	18079
GEAVEFSFPLAFTVE	19756
GEQAEFSFPLAFTVE	13000
GEQVAFSFPLAFTVE	13930
GEQVEASFPLAFTVE	6533
GEQVEFAFPLAFTVE	20072
GEQVEFSAPLAFTVE	7378
GEQVEFSFALRAFTVE	19480
GEQVEFSFPAAFTVE	10589

GEQVEFSFPLTFTVE	18318
GEQVEFSFPLAATVE	9572
GEQVEFSFPLAFAVE	19516
GEQVEFSFPLAFTAE	16765
GEQVEFSFPLAFTVA	18187
GEQVEFSFPLAFTVE	18219
ZZZZZZDTYICEVED	5017
ZZZZZZDTYICEVEZ	5421
ZZZZZZDTYICEVZZ	2166
ZZZZZZDTYICEZZZ	922
ZZZZZZDTYIZZZZ	564
ZZZZZZDTYICEVED	3031
EEVQLLVFGLTANS	23357
EEVQLLVFGLTANZ	15808
EEVQLLVFGLTANZ	16496
EEVQLLVFGLTAZZ	14097
EEVQLLVFGLTZZZ	16473
EEVQLLVFGLZ	10516
EEVQLLVFGZ	10372
EEVQLLVFZ	7333
EEVQLLVZ	1098
ZEVQLLVFGLTANS	16716
ZZVQLLVFGLTANS	5281
ZZZQLLVFGLTANS	4310
ZZZZLLVFGLTANS	1026
ZZZZZLVFGLTANS	664
ZZZZZZVFGLTANS	779
ZZZZZZZFGLTANS	760
ZZZZZZZZGLTANS	657
EEVQLLVFGLTANS	18040
THLLQGQSLTLTLES	10850
THLLQGQSLTLTLEZ	10269
THLLQGQSLTLTZZ	4668
THLLQGQSLTLTZZ	908
THLLQGQSLTLZZZ	844
THLLQGQSLTZ	475
THLLQGQSLZ	548
THLLQGQSZ	570
THLLQGQZ	442
ZHLLQGQSLTLTLES	11445
ZZLLQGQSLTLTLES	11631
ZZZLQGQSLTLTLES	7993
ZZZZQGQSLTLTLES	6887
ZZZZZGQSLTLTLES	3305
ZZZZZZQSLTLTLES	4453

ZZZZZZZSLTLTLES	1086
ZZZZZZZZLTLTLES	1201
THLLQGQSLTLTLES	9756
GEQVEFSFPLAFTVE	18856
GEQVEFSFPLAFTVZ	16222
GEQVEFSFPLAFTZZ	12535
GEQVEFSFPLAFZZZ	11384
GEQVEFSFPLAZZZZ	5846
GEQVEFSFPLZZZZZ	4749
GEQVEFSFPZZZZZ	2208
GEQVEFSFZZZZZZ	3277
GEQVEFSZzzzzzzz	742
ZEQVEFSFPLAFTVE	19736
ZZQVEFSFPLAFTVE	18684
ZZZVEFSFPLAFTVE	12892
ZZZZEFSFPLAFTVE	12166
ZZZZZFSFPLAFTVE	2134
ZZZZZZSFPLAFTVE	1454
ZZZZZZZFPLAFTVE	1391
ZZZZZZZZPLAFTVE	1489
GEQVEFSFPLAFTVE	18867
empty (control)	580

Example 11

This example characterizes CD4 receptor sequences found to have HIV gp120 binding activity in screening tests.

- 5 Panel A displays information obtained from sequential replacement of amino acid residues by alaninyl residues. In panel A, a (+) signifies a decrease in binding affinity whereas a (>) indicates that replacement of the residue by an alaninyl residue yields an increase in binding affinity. Sequences are shown with amino-terminus at the top and the carboxyl-terminus at the bottom. Right and left sides are from independent assays.
- 10

## 15 Panel A.

105-113	116-130	131-145	216-229
D	E	T	G

T	E	H	E
<u>++Y++</u>	V	L	Q
+I+	<u>+Q+</u>	L	+V+
C	<u>+L+</u>	<u>+Q+</u>	+E+
+E+	<u>+L+</u>	G	<u>++F++</u>
+V+	<u>+V+</u>	<u>+Q+</u>	S
+E+	<u>+F+</u>	S	<u>++F++</u>
D	G	<u>+L+</u>	P
	<u>+L</u>	T	<u>++L++</u>
	T	<u>+L++</u>	A
	A	>T>	<u>++F++</u>
N		<u>+++L+++</u>	T
S		<u>++E++</u>	V
D		S	E

Panel B indicates the effect on binding affinity when successive amino acid residues are deleted, either from  
 5 the amino-terminus (right side-symbols) or the carboxyl-terminus from the bottom (left side-symbol). A (+) signifies a decrease in binding affinity, and the underlined residues indicate which residue was the last residue to be serially deleted.

10

## Panel B.

105-113	116-130	131-145	216-229
<u>D</u> +	E	T	G
T	<u>E</u> +	H	E
<u>Y</u>	V+	<u>L</u> +	<u>Q</u> +
I	<u>Q</u> ++	L+	V+
<u>C</u>	<u>L</u> +++	<u>Q</u> ++	<u>E</u> +++
<u>++E</u>	<u>L</u> +++	G++	F+++
<u>++V</u>	V+++	<u>Q</u> +++	S++++
<u>+E</u>	<u>++++F</u> ++++	<u>+++S</u> +++	<u>++++F</u> ++++
D	<u>++G</u>	<u>+++L</u>	<u>+++P</u>
	+L	<u>+++T</u>	<u>+++L</u>
	T	<u>+++L</u>	<u>++A</u>
	A	<u>++T</u>	<u>++F</u>
N		<u>++L</u>	<u>+T</u>
S		<u>+E</u>	<u>+V</u>
D		S	E

All publications cited herein are hereby incorporated by reference to the same extent as if each publication were individually and specifically indicated to be incorporated by reference and were set forth in its 5 entirety herein.

While this invention has been described with an emphasis upon preferred embodiments, it will be obvious to those of ordinary skill in the art that variations of the preferred embodiments can be used and that it is 10 intended that the invention can be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications encompassed within the spirit and scope of the invention as defined by the following claims.

TOP SECRET - DEFENSE